

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2004, 14:37:17 ; Search time 39 Seconds
(without alignments)
2281.439 Million cell updates/sec

Title: PCT-US03-12556-2
Perfect score: 1522
Sequence: 1 MKKKVLAAMVALAAPVQS.....BAQRTAGYIHLWFTYYNR 282

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_25:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1257.5	82.6	283	052864	O52864 bacillus th
2	1257.5	82.6	283	Q81HW1	Q81HW1 bacillus ce
3	1251.5	82.2	283	Q81V27	Q81V27 bacillus an
4	1250.5	82.2	283	Q84FI6	Q84FI6 bacillus ce
5	482.5	31.7	275	Q84DK1	Q84DK1 listeria se
6	448	29.4	242	Q83TL2	Q83TL2 listeria se
7	442	29.0	242	Q84DK4	Q84DK4 listeria mo
8	442	29.0	242	Q83U53	Q83U53 listeria mo
9	440	28.9	242	Q84DK0	Q84DK0 listeria mo
10	437	28.7	242	Q84DK5	Q84DK5 listeria mo
11	437	28.7	242	Q84DK3	Q84DK3 listeria mo
12	411	27.0	236	Q84DK6	Q84DK6 listeria mo
13	411	27.0	242	Q84DK2	Q84DK2 listeria mo
14	357.5	23.5	223	Q84DK7	Q84DK7 listeria mo
15	325	21.4	154	Q8VMY6	Q8VMY6 listeria iv
16	314.5	20.7	178	Q48580	Q48580 listeria iv

17	230	15.1	398	2	Q8GCV3	Q8GCV3 clostridium
18	219	14.4	370	2	Q83Z25	Q83Z25 clostridium
19	166	10.9	272	16	Q97TF6	Q97TF6 clostridium
20	146	9.6	292	17	Q8PY92	Q8PY92 methanosarc
21	123	8.1	393	17	Q8TZ39	Q8TZ39 methanopyru
22	117.5	7.7	242	16	Q896L3	Q896L3 clostridium
23	111	7.3	1042	13	Q9DEG5	Q9DEG5 gallus gall
24	111	7.3	2402	13	Q9DEH4	Q9DEH4 gallus gall
25	109.5	7.2	245	16	Q97KA0	Q97KA0 clostridium
26	108.5	7.1	2140	16	Q97RY6	Q97RY6 streptococc
27	107.5	7.1	325	17	Q8TWV6	Q8TWV6 methanopyru
28	107	7.0	2144	2	Q8S4M8	Q8S4M8 streptococc
29	107	7.0	2144	16	Q8DQF7	Q8DQF7 streptococc
30	106.5	7.0	904	13	Q9DEG4	Q9DEG4 gallus gall
31	106	7.0	607	5	Q8ICW0	Q8ICW0 plasmodium
32	105	6.9	4144	6	Q8WN22	Q8WN22 canis famil
33	103	6.8	625	16	Q97J52	Q97J52 clostridium
34	102.5	6.7	470	16	Q8YX86	Q8YX86 anabaena sp
35	102.5	6.7	956	5	Q97Z38	Q97Z38 plasmodium
36	102	6.7	1818	16	Q8EVF1	Q8EVF1 mycoplasma
37	101.5	6.6	570	10	Q24041	Q24041 liriiodendro
38	100.5	6.6	477	5	Q25267	Q25267 leishmania
39	100.5	6.6	1740	5	Q81JK9	Q81JK9 plasmodium
40	100	6.6	957	16	Q8EWU3	Q8EWU3 mycoplasma
41	99	6.5	273	16	Q9PHU3	Q9PHU3 xylella fas
42	98.5	6.5	378	5	Q9GNZ4	Q9GNZ4 leishmania
43	98	6.4	352	16	Q88AF4	Q88AF4 pseudomonas
44	98	6.4	472	16	Q8CVP1	Q8CVP1 streptococc
45	98	6.4	988	15	Q8A3K2	Q8A3K2 bacteroides

ALIGNMENTS

RESULT 1

OS2864	PRELIMINARY;	PRT;	283 AA.
AC	O52864		
DT	01-JUN-1998	(Tremblrel. 06, Created)	
DT	01-JUN-1998	(Tremblrel. 06, Last sequence update)	
DT	01-OCT-2003	(Tremblrel. 25, Last annotation update)	
DE	Phosphatidyl-degrading phospholipase C.		
GN	PLCA.		
OS	Bacillus thuringiensis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=1428;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=Bt13;		
RX	PubMed=9732531;		
RA	Loevgren A., Carlson C.R., Eskils K., Kolsto A.B.;		
RT	"Localization of putative virulence genes on a physical map of the		
RT	Bacillus thuringiensis subsp. geatchiae chromosome.";		
RL	Curr. Microbiol. 37:245-250(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bt13;		
RA	Loevgren A.;		
RL	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; Y16268; CAA76148.1; -.		
DR	HSSP; P09598; 1AH7.		
DR	GO; GO:0004629; F:phospholipase C activity; IEA.		
DR	GO; GO:0008270; F:zinc ion binding; IEA.		
DR	InterPro; IPR008947; PLC_Nuclease.		
DR	InterPro; IPR001531; Zn_dep_PLPC.		
DR	Ffam; PF00882; Zn_dep_PLPC; 1.		
DR	PRINTS; P00479; PREPHLPASEC.		
DR	ProDom; PD003946; Zn_dep_PLPC; 1.		
DR	PROSITE; PS00384; PROKAR_ZN_DEPEND_PLPC; 1.		
SQ	SEQUENCE 283 AA; 32387 MW; 4D72E572F8A68FF CRC64;		

Query Match 82.6%; Score 1257.5; DB 2; Length 283;
Best Local Similarity 80.2%; Pred. No. 3.6e-90;

Matches	Conservative	Mismatches	Indels	Gaps
Qy	1 MKKKVLALAAWVLAAPQSVVPFACTNNSSPAIL-RWSAEDKHNGGINSHLWIVNRAI	59		
Dd	1 MKKKVLALAAAIITLVLPQSVAFAHENDGGORFGVI PRMSAEDKHKEGVNSHLWIVNRAI	60		
Qy	60 DIMSRNTTIYNPNETALLNEWRADELNGIYSADYENPYDNSTVASHPYDPDTGTITIPP	119		
Dd	61 DIMSRNTTLVKQRVALLNWRTELENGIYAADYENPYDNSTFASHYPDPDNGKTYIPY	120		
Qy	120 AKIAKETGAKYFNLAGOAYQNQMCOQAFYYLGLSLHYLDGVNQPMHAANFTNLSPMGFH	179		
Dd	121 AKQAKETGAKYFKLAGESYKNCKMKQAFYYLGLSLHYLDGVNQPMHAANFTNLSPQG FH	180		
Qy	180 SKYENFDVTIKNNYIVSDSGYNNWKGNABEDWTIEGAARAKODYPGVVNDTTDKWFVKA	239		
Dd	181 SKYENFDVTIKDNYKVTDGGYNNWKGNTPBEDWIHGAAAVAKQDIAGIVNDNTKWDFVRA	240		
Qy	240 AVSSEYADKNRAEYPTVTGKELMEAQRTAGYIHLWFDITYNVR	282		
Dd	241 AVSSEYADKNRAEYPTMTGKLMDAQRTAGYIOLWFEDTIGNR	283		

```

RESULT 2
Q81HW1
ID ID Q81HW1 PRELIMINARY; PRT; 283 AA.
AC Q81HW1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DE Phospholipase C (EC 3.1.4.3).
GN BC0670.
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP MEDLINE=22608415; PubMed=12721630;
RX Ivanova N., Scrokina A., Anderson I., Galleron N., Candelon B.,
RA Kapatal V., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Goltseman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.;
RA "Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis";
RL Nature 423:87-91(2003).
DR EMBL; AE017000; AAP07684.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004629; F:phospholipase C activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR008947; PLC_Nuclease.
DR InterPro; IPR001531; Zn_dep_PLP_C.
DR Pfam; PF00882; Zn_dep_PLPC_1.
DR PRINTS; PR00479; PRPHPLPASEC.
DR PRODOM; PRO03946; Zn_dep_PLPC; 1.
DR PROSITE; PS00384; PROKAR_ZN_DEPND_PLPC; 1.
DR HYDROLASE; Complete proteome.
SQ SEQUENCE 283 AA; 32387 MW; 4D72E6723F8A68FF CRC64;

```

Db	121	AKQAKGTGAKYFKLAGESYKNKDMKQAPFYVLGLSLHYLGDVNPQMAAANFTNLSYQGFPH	180
Qy	180	SKYENFVDITIKNNYIVSDNSGYNNWKGANPEDWIEGAAVAAKODYPGVNDTTKDMFVKA	239
Db	181	SKYENFVDITIKDNYKVYTDGNGYNNWKGTPEDWIHGAAVVAKODYAGIVNDTKDMFVRA	240
Qy	240	AVSQEYADKKRAEVTPTGKRLMEAQORVTAGYIHLWFDTYVNR	282
Db	241	AVSQEYADKKRAEVTPTGKRLMDAQORVTAGYIQLWFDTYGNR	283
RESULT 3			
Q81V27			
ID	Q81V27	PRELIMINARY;	PRT; 283 AA.
AC	Q81V27;		
DT	01-JUN-2003	(TrEMBLrel. 24, Created)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)	
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)	
DE	Phospholipase C.		
GN	PLC OR BA0677.		
OS	Bacillus anthracis (strain Ames).		

```

RESULT 3
Q81V27
ID ID PRELIMINARY; PRT; 283 AA.
AC Q81V27
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE DE Phospholipase C.
DE PLC OR BA0677.
GN GN
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX OX NCBI_TaxId=198094;
[1]
FN SEQUENCE FROM N.A.
RX MEDLINE=22608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Riststone J., Wu M.,
RA Kolony J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin H.M., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R., Redmond C., Theaite J.E., White O., Salzberg S.L.,
RA Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
RA Fraser C.M.;
RA "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
DR EMBL; AE017026; AAP24690.1; -.
DR TIGR; BA0677; -.
DR GO; GO:0004629; P:phospholipase C activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR008947; PLC Nuclease.
DR InterPro; IPR001531; Zn dep PLPC.
DR Pfam; PF00882; Zn dep PLPC; 1.
DR PRINTS; PR00479; PRPHPLPASEC.
DR PRODOM; PD003946; Zn dep PLPC; 1.
DR PROSITE; PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
DR Complete proteome.
SC SEQUENCE 283 AA: 32384 MW: 60C7296855D468A7 CRC64;

```

Query Match	82.24;	Score	1251.5;	DB	16;	Length	283;
Best Local Similarity	80.24;	Pred. No.	1.1e-89;				
Matches	227;	Conservative	23;	Mismatches	32;	Indels	1;
Gaps	17;						
Qy	1	MKKKVLALAAAMVALAAPQSVVFPACTVNNSESPAPIL-RWSAEDKHNEGINSHLWVNR	AI	59			
Db	1	MKKKVLALAAAIITVLAPLQNVAFAHENDGGSKIIVHRWSAEDKHGKVNSHLWVNR	AI	60			
Qy	60	DIMSRNTTIVNPNETALINWEARDLENGISYADYENPYDNTSYASHFPDPTDGT	YIPF	119			
Db	61	DIMSRNTTIVKQDRVAQLNWRTELENGIYAADYENPYDNTSPASHFPDPTDNGK	YIPF	120			
Qy	120	AKHAKETGAKYFNNLAGQAYQNQDMQQAFFYGLSLHYLGDVNPQMPHAAFN	NLSVP	179			
Db	121	AKQAKETGAKYFKLAGESYKKNKMQAFFYGLSLHYLGDVNPQMPHAAFN	NLSYP	180			
Qy	180	SKYENFVDITIKNNYITVSDSGYNNWKGANPEDWTGAAVAAKQDYPGVVND	ITTKDM	239			
Db	181	SKYENFVDITIKNNKVTVDGNGYNNWKGINPEDWTHGAAVVAKQDYS	GVINDNT	240			

QY 240 AVSQYADKRAEAVTPVTGKRLMEQORVTAGVYHILWFDTYVNR 282
 Db 241 AVSQYADKRAEAVTPVTGKRLMDAQRVTAGVYQLWFDTYGR 283

RESULT 4

ID Q84F16 PRELIMINARY; PRT; 283 AA.
 AC Q84F16;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Phosphatidylcholine-specific phospholipase C.
 GN PC-PLC.
 OS *Bacillus cereus*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1396;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=569;
 RA Pomerantsev A.P., Kalnin K.V., Leppla S.H.;
 RT "Bacillus cereus pc-plc and sph genes.";
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY195600; AAO40751.1; --
 DR GO: GO:0004629; F:phospholipase C activity; IEA.
 DR GO: GO:0008270; F:zinc ion binding; IEA.
 DR InterPro: IPR008947; PLC Nuclease.
 DR Pfam: PF00882; Zn dep PLPC_1.
 DR PRINTS: PR00479; PRPHPLPASEC.
 DR PRODOM: PD003946; Zn dep PLPC; 1.
 DR PROSITE: PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
 SQ SEQUENCE 283 AA; 32450 MW; 289787200D5B9FC8 CRC64;

Query Match 82.2%; Score 1250.5; DB 2; Length 283;
 Best Local Similarity 79.9%; Pred. No. 1.3e-89;
 Matches 226; Conservative 24; Mismatches 32; Indels 1; Gaps 1;

QY 1 MKKKVLALAAVVAAPVQSVFAQTNNSEGPAPIL-RWSAEDKHNEGINSHLWVNR 59
 Db 1 MKKKVLALAAITLVAPLQNVAFAHENDGGSKIKIVHWSAEDKHNEGINSHLWVNR 60
 QY 60 DIMSRNTTIVNPNETALINEWRADLENGIYSDYENPYDNTSYASHFYDPTGTITYPP 119
 Db 61 DIMSRNKTIVQDRVALLNEWRTELENGIYAADYENPYDNTSYASHFYDPTGTITYPP 120
 QY 120 AKHAKETGAKYFNLAGQYQNDQQOAPFYGLSLHYLGDVNPMPHAANFTNLSYPMGFH 179
 Db 121 AKQAKETGAKYFKLAGESYKKNQKQAPFYGLSLHYLGDVNPMPHAANFTNLSYPMGFH 180
 QY 180 SKYENFVDTIKNNYIVSDNSGYNWKGANPEDWIEGAAVAAKQDYPGVVNDTTKDWFK 239
 Db 181 SKYENFVDTIKNNYIVTNGYVWKGTPEDWIEGAAVAAKQDYPGVVNDTTKDWFK 240
 QY 240 AVSQYADKRAEAVTPVTGKRLMEQORVTAGVYHILWFDTYVNR 282
 Db 241 AVSQYADKRAEAVTPVTGKRLMDAQRVTAGVYQLWFDTYGR 283

RESULT 5

ID Q84DK1 PRELIMINARY; PRT; 275 AA.
 AC Q84DK1;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Phospholipase (Fragment).
 GN PLCB.
 OS *Listeria seeligeri*.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1640;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=22535694; PubMed=12648840;
 RA Schmid M., Walcher M., Bubert A., Wagner M., Wagner M.,
 RA Schleifer K.-H.;
 RT "Nucleic acid-based, cultivation-independent detection of *Listeria*
 RT spp. and genotypes of *L. monocytogenes*.";
 RL FEMS Immunol. Med. Microbiol. 35:215-225(2003).
 DR EMBL: AY150839; AAO19486.1; --
 DR GO: GO:0004629; F:phospholipase C activity; IEA.
 DR GO: GO:0008270; F:zinc ion binding; IEA.
 DR InterPro: IPR008947; PLC Nuclease.
 DR Pfam: PF00882; Zn dep PLPC.
 DR PRINTS: PR00479; PRPHPLPASEC.
 DR PRODOM: PD003946; Zn dep PLPC; 1.
 DR PROSITE: PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
 FT NON TER 275 275
 SQ SEQUENCE 275 AA; 31213 MW; 87DA8752CF1BEC18 CRC64;
 Query Match 31.7%; Score 482.5; DB 2; Length 275;
 Best Local Similarity 36.9%; Pred. No. 1e-29;
 Matches 97; Conservative 54; Mismatches 101; Indels 11; Gaps 3;
 QY 9 AAMVALAAPVQSVFAQTNNSESPAP-----ILRWSAEDKHNEGINSHLWVNR 59
 Db 14 ASIVALSQADEASACGDSVDQIAPHDIQNKLPKLGWSAHPKSNETHLWLFNQAE 73
 QY 60 DIMSRNTTIVNPNETALINEWRADLENGIYSDYENPYDNTSYASHFYDPTGTITYPP 119
 Db 74 KILAKHTVGAQLDLVRELKYNKIAQIGFDADHKNPYYDKNTFLSHFYNPKTKTYIAG 133
 QY 120 AKHAKETGAKYFNLAGQYQNDQQOAPFYGLSLHYLGDVNPMPHAANFTNLSYPMGFH 179
 Db 134 PPNAKDTGTYFNISIEYQDGNFEKAFYNLGLAIHYTYDISQPMHANNFTALSHPVGYH 193
 QY 180 SKYENFVDTIKNNYIVS-DSNGYNWKGANPEDWIEGAAVAAKQDYPGVVNDTTKDWFK 238
 Db 194 CAYENYVDTFKQIFQASAESEAKW-FTDDVSEPHENAKRAQADYPKIVNTIINKSYIQ 252
 QY 239 AAVSQYADKRAEAVTPVTGKRL 261
 Db 253 GLSDSQDRTWKAVRAATGKRL 275
 RESULT 6
 ID Q83TL2 PRELIMINARY; PRT; 242 AA.
 AC Q83TL2;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Phospholipase (Fragment).
 GN PLCB.
 OS *Listeria monocytogenes*.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IMVW 1630;
 RX MEDLINE=22535694; PubMed=12648840;
 RA Schmid M., Walcher M., Bubert A., Wagner M., Wagner M.,
 RA Schleifer K.-H.;
 RT "Nucleic acid-based, cultivation-independent detection of *Listeria*
 RT spp. and genotypes of *L. monocytogenes*.";
 RL FEMS Immunol. Med. Microbiol. 35:215-225(2003).
 DR EMBL: AY150830; AAO19477.1; --
 DR EMBL: AY150840; AAO19487.1; --
 DR GO: GO:0004629; F:phospholipase C activity; IEA.
 DR GO: GO:0008270; F:zinc ion binding; IEA.
 DR InterPro: IPR008947; PLC Nuclease.
 DR Pfam: PF00882; Zn dep PLPC.
 DR PRINTS: PR00479; PRPHPLPASEC.
 DR PRODOM: PD003946; Zn dep PLPC; 1.
 DR ProDom; PD003946; ZN_dep_PLPC; 1.

DR PROSITE; PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
FT NON TER 242
SQ SEQUENCE 242 AA; 27963 MW; CA45226804B99E87 CRC64;

Query Match 29.4%; Score 448; DB 2; Length 242;
Best Local Similarity 39.4%; Pred. No. 4.4e-27;
Matches 93; Conservative 45; Mismatches 86; Indels 12; Gaps 4;

QY 2 KKKVLALAAVMA-LAAPVQSVV-----FAQT-----NNSESPAP-IIRWSAEKHNREGIN 49
DB 4 KKKVLGLMCLIASLVFPVTTIKANACCDBEYLQTPAAPHDIDSKLPKLSWSADNPNTNDVN 63

QY 50 SHLIVNRAIDIMSRNTTIVNPNETALLNEWRADLENGIYSADYENPYDNSTYASHFYD 109
DB 64 THYWLFKQAEKILAKOVNHRANLWELKFKDKQIAQGIYDADHKNPYDTSTFLSHFYN 123

QY 110 PDTGTYTIPFAKHAKETGAKYFNLAGQAYQNDMQQAFYYLGLSLHYLGDVNPQMHAANF 169
DB 124 PDRNTYILPGFANAKITGAKYFNQSVTDYREGKFDTPAFYKGLAIHYVTDISQPMHANF 183

QY 170 TNLSPYMGFHSKYENFVDTIKNNYIVSDSGYNNWKGANPEDWIEGRAVAAKQDYP 225
DB 184 TAISYPGPHYCAVENYVDTIKHNYQATEDMVAKRFSDDVDKDLNENAKRAKRDYP 239

RESULT 7
ID Q84DK4 PRELIMINARY; PRT; 242 AA.
AC Q84DK4;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Phospholipase (Fragment).
GN PLCB.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22535694; PubMed=12648840;
RA Schmid M., Walcher M., Bubert A., Wagner M., Wagner M.,
RA Schleifer K.-H.;
RT "Nucleic acid-based, cultivation-independent detection of Listeria
RT spp. and genotypes of L. monocytogenes."
RL FEMS Immunol. Med. Microbiol. 35:215-225(2003).
DR EMBL; AY150833; AAO19480.1; -.
DR GO; GO:0004629; F:phospholipase C activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR008947; PLC_Nuclease.
DR Pfam; PF00882; Zn_dep_PLPC; 1.
DR PRINTS; PR00479; PRPHPLPASEC.
DR ProDom; PD003946; Zn_dep_PLPC; 1.
DR PROSITE; PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
FT NON TER 242
SQ SEQUENCE 242 AA; 27892 MW; 91D9A192EBE4DE18 CRC64;

Query Match 29.0%; Score 442; DB 2; Length 242;
Best Local Similarity 42.6%; Pred. No. 1.3e-26;
Matches 81; Conservative 38; Mismatches 71; Indels 0; Gaps 0;

QY 36 LRWSAEKHNREGINSHLWVNRNRAIDIMSRNTTIVNPNETALLNEWRADLENGIYSADYEN 95
DB 50 LWSADNPNTDVTNHYLWFKQAEKILAKOVNHRANLWELKFKDKQIAQGIYDADHKN 109

QY 96 PYYDNSTYASHFYDPTGTYTIPFAKHAKETGAKYFNLAGQAYQNDMQQAFYYLGLSLH 155
DB 110 PYYDTSTFLSHFYNPDKNTYLPGFANAKITGAKYFNQSVADYREGKFDTPAFYKGLAIH 169

QY 156 YLGDVNPQMHAANFTNLSPYMGFHSKYENFVDTIKNNYIVSDSGYNNWKGANPEDWIEG 215
DB 170 YTTDISQPMHANFTAISYPGPHYCAVENYVDTIKHNYQATEDMVAKRFSDDVDKDLNENAKRAKRDYP 239

QY 216 AAVAAKQDYP 225
DB 230 NAKRAKRDYP 239

RESULT 9
ID Q84DK0 PRELIMINARY; PRT; 242 AA.
AC Q84DK0;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Phospholipase (Fragment).
GN PLCB.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2004, 14:32:07 ; Search time 13 Seconds
(without alignments)
1129.522 Million cell updates/sec

Title: PCT-US03-12556-2
Perfect score: 1522
Sequence: 1 MKKVLALAMVALAPVQS.....EAQRTAGYIHLWFDYVNR 282

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1253.5	82.4	283	1 PHLD_BACCE	P33376 bacillus ce
2	1252.5	82.3	283	1 PHLC_BACCE	P09598 bacillus ce
3	511	33.6	289	1 PHLC_LISMO	P33378 listeria mo
4	254	16.7	398	1 PHLC_CLOBI	P20419 clostridium
5	244.5	16.1	399	1 PHLC_CLOSO	Q8VUZ6 clostridium
6	237	15.6	398	1 PLC2_CLOPE	Q9RF12 clostridium
7	229	15.0	398	1 PLC1_CLOPE	P15310 clostridium
8	213.5	14.0	399	1 PHLC_CLOHA	P59026 clostridium
9	205	13.5	398	1 PHLC_CLONO	Q46150 clostridium
10	102	6.7	562	1 EXG2_YEAST	P52911 saccharomyc
11	102	6.7	2358	1 MOXD_SCHPO	Q9Y719 schizosacch
12	97.5	6.4	363	1 OMPF_SALTI	Q56113 salmonella
13	97.5	6.4	363	1 OMPF_SALTY	P37432 salmonella
14	97	6.4	468	1 YBFM_ECOLI	P75733 escherichia
15	95.5	6.3	1283	1 OSH2_YEAST	Q12451 saccharomyc
16	95	6.2	624	1 YE70_METJA	Q58865 methanococc
17	94.5	6.2	418	1 HLT_VIBPA	Q99289 vibrio para
18	94	6.2	623	1 EXAA_PSEAE	Q9Z4J7 pseudomonas
19	94	6.2	837	1 XYNZ_CLOTM	P10478 clostridium
20	93.5	6.1	277	1 FMF4_ECOLI	P11900 escherichia
21	93.5	6.1	369	1 XYNA_PPRERU	P48789 prevotella
22	93.5	6.1	787	1 OXAA_CHLMO	Q9PK63 chlamydia m
23	93	6.1	333	1 XLNA_AGABI	O60206 agaricus bi
24	92.5	6.1	910	1 DNMJ_MYCPN	P75354 mycoplasma
25	92	6.0	782	1 Y044_UREPA	Q9P2A1 ureaplasma
26	91.5	6.0	474	1 PEDA_LACHE	Q48558 lactobacill
27	91	6.0	377	1 CCHL_SCHPO	O74794 schizosacch
28	90.5	5.9	257	1 OMEV_VIBAN	O08237 vibrio angu
29	90.5	5.9	787	1 OXAA_CHLTR	O84253 chlamydia t
30	90	5.9	423	1 P65_MYCPN	P53663 mycoplasma
31	90	5.9	1365	1 GTF5_STRDO	P29336 streptococ
32	89	5.8	235	1 NCAP_BUNSH	P03513 bunyavirus
33	89	5.8	695	1 CICA_PAEPP	P57092 paenibacill

ALIGNMENTS

RESULT 1

ID	PHLD_BACCE	STANDARD;	PRT;	283 AA.
AC	P33376;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Phospholipase C precursor (EC 3.1.4.3) (PLC) (Phosphatidylcholine			
DE	cholinephosphohydrolase) (Cereolysin A).			
GN	CERA.			
OS	Bacillus cereus.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1396;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=GP-4;			
RC	MEDLINE=89121149; PubMed=2536680;			
RA	Gilmore M.S., Cruz-Rodz A.L., Lelmeister-Waechter M., Kreft J.,			
RA	Goebel W.;			
RT	"A Bacillus cereus cytolytic determinant, cereolysin AB, which			
RT	comprises the phospholipase C and sphingomyelinase genes: nucleotide			
RT	sequence and genetic linkage."			
RL	J. Bacteriol. 171:744-753(1989).			
RN	[2]			
RP	SEQUENCE OF 50-142 FROM N.A.			
RA	Gilmore M.S., Gilmore K.S., Goebel W.;			
RA	"A new strategy for ordered DNA sequencing based on a novel method for			
RT	the rapid purification of near-milligram quantities of a cloned			
RT	restriction fragment."			
RL	Gene Anal. Tech. 2:108-114(1985).			
CC	- FUNCTION: Required, with sphingomyelinase to effect target cell			
CC	lysis (hemolysis).			
CC	- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-			
CC	diacylglycerol + choline phosphate.			
CC	- COFACTOR: Binds 3 zinc ions.			
CC	- SUBUNIT: Monomer.			
CC	- SIMILARITY: BELONGS TO THE BACTERIAL ZINC-METALLOPHOSPHOLIPASES C			
CC	FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M24149; AAA91819.1; -			
DR	EMBL; M35411; AAA22524.1; -			
DR	PIR; S18978; P50197.			
DR	HSSP; P09598; 1AH7.			
DR	InterPro; IPR008947; PLC Nuclease.			
DR	InterPro; IPR001531; Zn_dep_PLPC.			
DR	Pfam; PF00882; Zn_dep_PLPC_1.			
DR	PRINTS; PR00479; PRPHPLPASEC.			
DR	ProDom; PD003946; Zn_dep_PLPC; 1.			

34	89	5.8	866	1	YCBS_ECOLI	P75857 escherichia
35	89	5.8	1068	1	TRI_SULSO	P93871 sulfolobus
36	89.5	5.8	858	1	SYL_VIBCH	Q9KTE6 vibrio chol
37	88	5.8	472	1	PEDA_STRFY	Q9A0M0 streptococ
38	88	5.8	484	1	COAT_IRV9	O39163 wiseana iri
39	88	5.8	695	1	PLB2_CANGA	Q8T906 candida gla
40	88	5.8	1956	1	ATX1_PLAFA	Q04956 plasmodium
41	87.5	5.7	407	1	TRPB_STRRG	Q8DM8 streptococ
42	87	5.7	295	1	YF67_ARCFU	O28705 archaeoglob
43	87	5.7	368	1	YJHT_ECOLI	P39371 escherichia
44	87	5.7	857	1	SYL_VIBPA	Q87RQ0 vibrio para
45	87	5.7	857	1	SYL_VIBVU	Q8DFE2 vibrio vuln

```

DR PROSITE, PS00384; PROKAR_ZN_DEPEND PLPC; 1.
KW Hydrolase; Zinc; Signal; Zymogen; Hemolysis.
FT SIGNAL 1 24
FT PROPEP 25 38
FT CHAIN 39 283
FT METAL 39 39
FT METAL 52 52
FT METAL 93 93
FT METAL 107 107
FT METAL 156 156
FT METAL 160 160
FT METAL 166 166
FT METAL 180 180
FT METAL 184 184
SQ SEQUENCE 283 AA; 32373 MW; 24D2E753402A6A44 CRC64;

Query Match 82.4%; Score 1253.5; DB 1; Length 283;
Best Local Similarity 79.9%; Pred. No. 3.7e-92;
Matches 226; Conservative 23; Mismatches 33; Indels 1; Gaps 1;

QY 1 MKKKYLALAAAMVALAAPVQSVVFAQTNNSESPAPIL-RWSAEDKNEGINSHLWIVNRAI 59
Db 1 MKKKYLALGAATLVAPLQSAFAHENDGQRFVGPVPSAEDKHKEGVNSHLWIVNRAI 60
QY 60 DIMSRNTTVNPETALLNEWRADLENGIYSADYENPYDNTSYASHFYDPTGTITYPF 119
Db 61 DIMSRNTTVLQKDRVALLNEWRTELENGIYAADYENPYDNTSFASHFYDPTGTITYPF 120
QY 120 AKHAKETGAKYFNLAGQYQNDMQQAFYGLSLHYLDGVNQPMHAAFNLTSLYPMGFH 179
Db 121 AKQAKETGAKYFKLAGESYKNDKQAFYGLSLHYLDGVNQPMHAAFNLTSLYPMGFH 180
QY 180 SKYENFVDTIKNYIVSNGSNWKNKANGPEWISGAAVAADKQFPGVNDTIDKDFYKA 239
Db 181 SKYENFVDTIKNYIKYKTDGNGYNNKGTNPEDWIHGAAVAADKQFAGIYNDNTKDFYRA 240
QY 240 AVSQEYADKWRARVPTVKRLWEAQRTAGVIHLMFDTYVNR 282
Db 241 AVSQEYADKWRARVPTVKRLWEAQRTAGVIHLMFDTYVNR 283

RESULT 2
PHLC_BACCE STANDARD; PRT; 283 AA.
AC P03538
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospholipase C precursor (BC 3.1.4.3) (PLC) (Phosphatidylcholine
DE cholinephosphohydrolase) (Cereolysin A).
GN PLC.
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SE-1;
RX MEDLINE=88313678; PubMed=3137122;
RA Johansen T., Holm T., Gudal P.H., Sletten K., Haugli F.B., Little C.;
RT "Cloning and sequencing of the gene encoding the phosphatidylcholine-
RT preferring phospholipase C of Bacillus cereus.";
RL Gene 65:293-304(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VNM B-164;
RX MEDLINE=93249510; PubMed=8387306;
RA Kuzmin N.P., Gavrilenko I.V., Krukov V.M., Karpov A.V.;
RT "Nucleotide sequence of phospholipase C and sphingomyelinase genes
RT from Bacillus cereus BKM-B164 (letter).";
RL Bioorg. Khim. 19:133-138(1993).
RN [3]
RP SEQUENCE OF 166-283 FROM N.A.
RC STRAIN=IAM 1208;

```

```

RX MEDLINE=88296483; PubMed=2841128;
RA Yamada A., Tsukagoshi N., Uda S., Sasaki T., Makino S., Nakamura S.,
RA Little C., Tomita M., Ikezawa H.;
RT "Nucleotide sequence and expression in Escherichia coli of the gene
RT coding for sphingomyelinase of Bacillus cereus.";
RL Eur. J. Biochem. 175:213-220(1988).
RN [4]
RP SEQUENCE OF 39-65.
RX MEDLINE=78043154; PubMed=72664;
RA Otnaess A.-B., Little C., Sletten K., Wallin R., Johnsen S.,
RA Fløengrud R., Prydz H.;
RT "Some characteristics of phospholipase C from Bacillus cereus.";
RL Eur. J. Biochem. 79:459-468(1977).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE=89159438; PubMed=2493587;
RA Hough E., Hansen L.K., Birkenes B., Jynge K., Hansen S., Hordvik A.,
RA Little C., Dodson E., Derewenda Z.;
RT "High-resolution (1.5 Å) crystal structure of phospholipase C from
RT Bacillus cereus.";
RL Nature 338:357-360(1989).
CC -1- FUNCTION: Required, with sphingomyelinase to effect target cell
CC lysis (hemolysis).
CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
CC diacylglycerol + choline phosphate.
CC -1- COFACTOR: Binds 3 zinc ions.
CC -1- SUBUNIT: Monomer.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL ZINC-METALLOPHOSPHOLIPASES C
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; X64141; CAA45502.1; -
CC EMBL; X12854; CAA31332.1; -
CC EMBL; X12711; CAA31213.1; -
CC EMBL; X64140; CAA45501.1; ALT_TERM.
CC PIR; S18978; PS0197.
CC PDB; 1AH7; 10-DEC-97.
CC InterPro; IPR008947; PLC Nuclease.
CC InterPro; IPR001531; Zn dep PLPC.
CC Pfam; PF00882; Zn dep PLPC; 1.
CC PRINTS; PR00479; PRPHPLPASEC.
CC ProDom; PD003946; Zn dep PLPC; 1.
CC PROSITE; PS00384; PROKAR_ZN_DEPEND PLPC; 1.
KW Hydrolase; Zinc; Signal; Zymogen; Hemolysis; 3D-structure.
FT SIGNAL 1 24
FT PROPEP 25 38
FT CHAIN 39 283
FT METAL 39 39
FT METAL 52 52
FT METAL 93 93
FT METAL 107 107
FT METAL 156 156
FT METAL 160 160
FT METAL 166 166
FT METAL 180 180
FT METAL 184 184
FT VARIANT 212 212
FT VARIANT 226 226
FT VARIANT 239 239
FT VARIANT 282 282
FT HELIX 47 49
FT HELIX 51 65
FT HELIX 72 80
FT TURN 81 81
FT HELIX 82 91
FT TURN 92 94

```

```

FT TURN 96 103
FT STRAND 105 107
FT TURN 110 110
FT TURN 111 114
FT STRAND 115 115
FT TURN 119 120
FT HELIX 124 140
FT TURN 141 142
FT HELIX 144 161
FT TURN 162 162
FT TURN 164 169
FT TURN 172 173
FT TURN 177 178
FT HELIX 179 190
FT TURN 191 194
FT TURN 203 204
FT HELIX 210 223
FT TURN 224 224
FT HELIX 225 227
FT TURN 228 228
FT HELIX 231 240
FT TURN 241 242
FT HELIX 244 280
SQ SEQUENCE 283 AA; 32363 MW; ACS452EFP2E22B19 CRC64;

Query Match 82.3%; Score 1252.5; DB 1; Length 283;
Best Local Similarity 80.2%; Pred. No. 4.4e-92;
Matches 227; Conservative 23; Mismatches 32; Indels 1; Gaps 1;

Oy 1 MKKKVLAAALAAAPVQSVVFAQTNNSSEAPIL-RWSAEDKNEGINSHLWVNR 59
Db 1 MKKKVLAAALAAITVAPLQSVAFAHENDGGSKIKIVHRWSAEDKHEGVNSHLWVNR 60

Oy 60 DIMSNTTIVNPNETALLNEWRADLENGIYSADYENPPYDNTSYASHFYDPDTGTYIPF 119
Db 61 DIMSNTTIVKQDRVAQLNEWTLENGIYAADYENPPYDNTSFASHFYDPDNGKTYIPF 120

Oy 120 AKHAKETGAKYFNLAGQYQNMOMQAFYLGSLHYLDGVNQPMHAAFNLTLSYPMGFH 179
Db 121 AKQAKETGAKYFKLAGESYKNDKMQAFYLGSLHYLDGVNQPMHAAFNLTLSYPMGFH 180

Oy 180 SKYENFVDTIKNNYTVSDSGYNNKGANPEWIEGAAVAAKODYPGVVNDTKDMFVKA 239
Db 181 SKYENFVDTIKNNYTVSDSGYNNKGANPEWIEGAAVAAKODYSYGVVNDTKDMFVKA 240

Oy 240 AVSQEYADKRAEVPVTKRLMEARVTAGYIHLWFDYVNR 282
Db 241 AVSQEYADKRAEVPVTKRLMDAQRVTAGYIQLWFDYVNR 283

RESULT 3
PHLC LISMO STANDARD; PRT; 289 AA.
ID PHLC LISMO STANDARD; PRT; 289 AA.
AC P33378;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospholipase C precursor (EC 3.1.4.3) (PLC) (Phosphatidylcholine
DE cholinephosphohydrolase) (Lecithinase).
GN PLCB OR PRTC OR LMO0205.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LO28 / Serovar 1/2c;
RX MEDLINE=92104678; PubMed=1309513;
RA Vazquez-Boland J.-A., Kocks C., Dramsi S., Ohayon H., Geoffroy C.,
RA Mengaud J., Cossart P.;
RT "Nucleotide sequence of the lecithinase operon of Listeria
RT monocytogenes and possible role of lecithinase in cell-to-cell
RT spread.";
RL Infect. Immun. 60:219-230(1992).

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=EGD / Serovar 1/2a;
RX MEDLINE=92258410; PubMed=1582425;
RA Domann E., Wehlant J., Rohde M., Pistor S., Hartl M., Goebel W.,
RA Leimeister-Waechter M., Wuensther M., Chakraborty T.;
RT "A novel bacterial virulence gene in Listeria monocytogenes required
RT proline-rich region of vinculin.";
RL EMBO J. 11:1981-1990(1992).

RN RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deboux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunat F., Kurapkut G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Rammel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlant J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).

RN RP SEQUENCE OF 144-152.
RX MEDLINE=91267617; PubMed=1904842;
RA Geoffroy C., Raveneau J., Beretti J.L., Lecroisey A.,
RA Vazquez-Boland J.-A., Alouf J.E., Berche P.;
RT "Purification and characterization of an extracellular 29-kilodalton
RT phospholipase C from Listeria monocytogenes.";
RL Infect. Immun. 59:2382-2388(1991).

CC -!- FUNCTION: IMPORTANT ROLE IN THE INFECTIOUS PROCESS. MAY
CC CONTRIBUTE TO EFFICIENT LYSIS OF THE TWO-MEMBRANE VACUOLES THAT
CC SURROUND THE BACTERIA AFTER DIRECT CELL-TO-CELL SPREAD.
CC -!- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
CC diacylglycerol + choline phosphate.
CC -!- COFACTOR: Binds 3 zinc ions (By similarity).
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL ZINC-METALLOPHOSPHOLIPASES C
CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; M82881; AAA25270.1; -
CC EMBL; X59723; CA42408.1; -
CC EMBL; AL591974; CAD00732.1; -
CC PIR; AF1100; AF1100.
CC PIR; S20888; S20888.
CC HSSP; P09598; 1A87.
CC Listlist; LMO00205; -
CC InterPro; IPR008947; PLC Nuclease.
CC InterPro; IPR001531; Zn_dep_PLPC.
CC Pfam; PF00882; Zn_dep_PLPC; 1.
CC PRINTS; PR00479; PRPHPLPASEC.
CC ProDom; PD03946; Zn_dep_PLPC; 1.
CC PROSITE; PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
KW Hydrolase; Zinc; Zymogen; Signal; Complete proteome.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 51 BY SIMILARITY.
FT CHAIN 52 289 PHOSPHOLIPASE C.
FT METAL 52 52 ZINC 3 (BY SIMILARITY).
FT METAL 65 65 ZINC 3 (BY SIMILARITY).
FT METAL 106 106 ZINC 1 (BY SIMILARITY).
FT METAL 120 120 ZINC 1 (BY SIMILARITY).

```

FT METAL 169 169 ZINC 1 (BY SIMILARITY).
FT METAL 173 173 ZINC 1 AND 3 (BY SIMILARITY).
FT METAL 179 179 ZINC 2 (BY SIMILARITY).
FT METAL 193 193 ZINC 2 (BY SIMILARITY).
FT METAL 197 197 ZINC 2 (BY SIMILARITY).
FT CONFLICT 5 5 K -> N (IN REF. 2).
FT CONFLICT 13 13 I -> T (IN REF. 2).
FT CONFLICT 222 222 D -> G (IN REF. 2).
SQ SEQUENCE 289 AA; 33277 MW; 0F35A2A3EDA6E372 CRC64;

Query Match 33.6%; Score 511; DB 1; Length 289;
Best Local Similarity 37.1%; Pred. NO. 2.2e-33;
Matches 106; Conservative 55; Mismatches 107; Indels 18; Gaps 5;

QY 2 KKKVLALAAVA-LAAPQSVV-----PAQT-----NNSEGPAP-ILWMSAEDKHNEGIN 49
DB 4 KKKVLGMLIASLVLPVPTIRKANACCBYLOTFAAPHDIDSKLPKHSWADNPTNTDVN 63

QY 50 SHLVNVRADIMSRNTTIVNPNTALLNWRADLENGIYADYENPYDNTSYASHFYD 109
DB 64 THYMLFKQAEKILAKDVNHRANLWNLKKPKDKQIAQGIYDADHKNPYDSTFLSHFYN 123

QY 110 PDTGTYTIPFAKAKETGAKYFNLAGQAYQNDMQQAFYGLSLHYHLYGDNVQPMHAANF 169
DB 124 PDRDNTYLPFGANAKITGAKYFNQSVTDYREGKFDYAFYKGLAIHYTYDISQPMHANF 183

QY 170 TNLSPMGFHSKYENFVDTIKONYIVSDSNGYWNWKGANPEDWIEGAVAQAQDYPGVVN 229
DB 184 TAISYPPGYHCAYENYVDTIRKHYQATSDMVAKRFCSDDVDKWLNYENAKRADYPKLVN 243

QY 230 DTTKDPFKAAVQSYADKRAEVTPTGKRLMEARQVTAQYIHLW 275
DB 244 AKTKSYLVG-----NSEWKDTVEPTGARLDSQQTLAGLEFW 283

RESULT 4
PHLC_CLOBI
ID PHLC_CLOBI STANDARD; PRT; 398 AA.
AC P20A19; Q8VUZA; Q9S532;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phospholipase C precursor (EC 3.1.4.3) (PLC) (Phosphatidylcholine
phosphatidylcholine) (Cbp).
GN PLC OR Cbp.
OS Clostridium bifermentans.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1490;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=ATCC 638;
RX MEDLINE=89108588; PubMed=2536356;
RA Tao J.Y., Siebel C.;
RT "Cloning and expression of the phospholipase C gene from Clostridium
perfringens and Clostridium bifermentans.";
RL Infect. Immun. 57:468-476(1989).
RN [2]
RP SEQUENCE FROM N.A., AND CONSTRUCTION OF A HYBRID ENZYME.
RC STRAIN=ATCC 638;
RX MEDLINE=99307183; PubMed=10377104;
RA Jepson M., Howelle A.M., Bullifent H.L., Bolgiano B., Crane D.T.,
RA Miller J., Holley J., Jayasekera P., Titball R.W.;
RT "Differences in the carboxy-terminal (putative phospholipid binding)
domains of Clostridium perfringens and Clostridium bifermentans
phospholipases C influence the hemolytic and lethal properties of
these enzymes.";
RL Infect. Immun. 67:3297-3301(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1012;
RA Karasawa T., Wang X., Maegawa T., Michiwa Y., Miwa K., Nakamura S.;
RT "C. sordellii phospholipase C.";

Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
[4]
RN REVIEW
RX MEDLINE=20465460; PubMed=11008117;
RA Jepson M., Titball R.W.;
RT "Structure and function of clostridial phospholipases C.";
RL Microbes Infect. 2:1277-1284(2000).
CC -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
cell membranes and cause cell rupture. Binds to eukaryotic
membranes where it hydrolyzes phosphatidylcholine. This enzyme has
10-fold less activity towards sphingomyelin than its C.perfringens
counterpart, is approximately 100-fold less hemolytic against
mouse erythrocytes and at least 100-fold less toxic in mice.
CC -!- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
diacylglycerol + choline phosphate.
CC -!- COPACTOR: Binds 3 zinc and 3 calcium ions (By similarity).
CC -!- DOMAIN: The protein is composed of 2 domains; the N-terminal
domain contains the phospholipase C active site (PLC), in a cleft
which is also occupied by the 3 zinc ions. The C-terminal domain
is a putative phospholipid-recognition domain, which shows
structural homology with phospholipid-binding C2-like domains from
a range of eukaryotic proteins. The ability to bind membrane
phospholipids in a Ca(2+) dependent manner is conferred by this C-
terminal domain (By similarity).
CC -!- MISCELLANEOUS: A hybrid protein between the N-terminus of
C.bifermentans and the C-terminus of C.perfringens has an activity
intermediate between the two.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL ZINC-METALLOPHOSPHOLIPASES C
FAMILY.
CC -!- SIMILARITY: Contains 1 PLAT domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC EMBL; AF072123; AAD41623.1; -.
DR EMBL; AB061869; BAB83265.1; -.
PIR; B30565; B30565.
DR InterPro; IPR001024; Lipoxigenase_LH2.
DR InterPro; IPR008976; PLAT_LH2.
DR InterPro; IPR008947; PLC_Nuclease.
DR InterPro; IPR001531; Zn_dep_PLPC.
DR Pfam; PF01477; PLAT; 1.
DR Pfam; PF00882; Zn_dep_PLPC; 1.
DR PRINTS; PR00479; PRPHPLPASEC.
DR PRODOM; PD003946; Zn_dep_PLPC; 1.
DR PROSITE; PS50095; PLAT; 1.
DR PROSITE; PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
KW Hydrolase; Toxin; Hemolysis; Calcium-binding; Zinc;
KW Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 398 PHOSPHOLIPASE C.
FT DOMAIN 29 272 PHOSPHOLIPASE C.
FT DOMAIN 273 281 LINKER.
FT DOMAIN 282 398 PLAT.
FT METAL 27 27 ZINC 1 (BY SIMILARITY).
FT METAL 37 37 ZINC 1 (BY SIMILARITY).
FT METAL 82 82 ZINC 2 (BY SIMILARITY).
FT METAL 94 94 ZINC 3 (BY SIMILARITY).
FT METAL 152 152 ZINC 3 (BY SIMILARITY).
FT METAL 156 156 ZINC 1 AND 3 (BY SIMILARITY).
FT METAL 162 162 ZINC 2 (BY SIMILARITY).
FT METAL 174 174 ZINC 2 (BY SIMILARITY).
FT METAL 178 178 ZINC 2 (BY SIMILARITY).
FT METAL 297 297 CALCIUM 1 (VIA CARBONYL OXYGEN) (BY
SIMILARITY).
FT METAL 298 298 CALCIUM 3 (BY SIMILARITY).
FT METAL 299 299 CALCIUM 3 (BY SIMILARITY).

QY 93 YENPYDINSTYASHFYDPTGTYYI-----PFAXHAKETGAKYFNLAGQAYQNDM 143
Db 85 YDPNAYD--LFQDHFDDPTGNFYLNDKNVYVAGPYDTASTQVTKFATLAEWKKGN 142
QY 144 QOAFYFLGLSLHYLDGVNQPMAHNFNTLSYPMGFHSHYENFVDTIKNNYIVSDSNGYWN 203
Db 143 KEATFLLQGMHYLDGLNTPYHAANVTAVDSP--GHVKYETFVENRKENYALNTAG--- 196
QY 204 WGANPEWISGAAVAQKQDPGVVNDTKWFFVKAASQYVADKWRAREVTPVTKRLME 263
Db 197 -----NDTTOGIY-KDAVANKDFQW-----MTQNSVKY 224
QY 264 AQRVTAGYI-----HLWFD 277
Db 225 AKKAKALYISHSTMKHND 244

RESULT 6

PLC2_CLOPE
ID PLC2_CLOPE STANDARD; PRT; 398 AA.
AC Q9RF12,
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phospholipase C precursor (EC 3.1.4.3) (PLC) (Phosphatidylcholine
cholinephosphohydrolase) (Alpha-toxin) (Hemolysin) (Lecithinase).
PLC OR CPA.
GN Clostridium perfringens.
OS Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1502;
RN [1] SEQUENCE FROM N.A., CHARACTERIZATION, AND X-RAY CRYSTALLOGRAPHY (2.4
RP ANGSTROMS) OF THE OPEN FORM.
RC STRAIN=SWCP;
RX MEDLINE=22005774; PubMed=12009886;
RA Justin N., Walker N., Bullifant H.L., Songer G., Bueschel D.M.,
RA Jost H., Naylor C.E., Miller J., Moss D.S., Tibball R.W., Basak A.K.;
RT "The first strain of Clostridium perfringens isolated from an avian
RT source has an alpha-toxin with divergent structural and kinetic
RT properties.";
RL Biochemistry 41:6253-6262 (2002).
RN [2]
RN [3]
RA Titball R.W., Naylor C.E., Basak A.K.;
RT "The Clostridium perfringens alpha-toxin.";
RL Anaerobe 5:51-64 (1999).
RN [3]
RN REVIEW.
RX MEDLINE=20465460; PubMed=11008117;
RA Japeon M., Tibball R.W.;
RT "Structure and function of clostridial phospholipases C.";
RL Microbes Infect. 2:1277-1284 (2000).
CC -1- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC cell membranes and cause cell rupture. Constitutes an essential
CC virulence factor in gas gangrene. Binds to eukaryotic membranes
CC where it hydrolyzes both phosphatidylcholine and sphingomyelin,
CC causing cell rupture. The diacylglycerol produced can activate
CC both the arachidonic acid pathway, leading to modulation of the
CC inflammatory response cascade and thrombosis, and protein kinase
CC C, leading to activation of eukaryotic phospholipases and further
CC membrane damage.
CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
CC diacylglycerol + choline phosphate.
CC -1- COFACTOR: Binds 3 zinc and 3 calcium ions.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: The protein is composed of 2 domains; the N-terminal
CC domain contains the phospholipase C active site (PLC), in a cleft
CC which is also occupied by the 3 zinc ions. The C-terminal domain
CC is a putative phospholipid-recognition domain, which shows
CC structural homology with phospholipid-binding C2-like domains from
CC a range of eukaryotic proteins. The ability to bind membrane
CC phospholipids in a Ca(2+) dependent manner and toxicity is

conferred by this C-terminal domain, which also contributes to the
sphingomyelinase activity.
-1- MISCELLANEOUS: This bacteria was isolated from a diseased swan.
-1- SIMILARITY: Contains 1 PLAT domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF204209; AAF20094.1; -.
DR PDB; 1KHO; 19-JUN-02.
DR InterPro; IPR001024; Lipoxigenase_LH2.
DR InterPro; IPR008976; PLAT_LH2.
DR InterPro; IPR008947; PLC_Nuclease.
DR InterPro; IPR001531; Zn_dep_PLPC.
DR Pfam; PF00882; Zn_dep_PLPC; 1.
DR PRINTS; PD00479; BRPHPLPASEC.
DR PRODOM; PD003946; Zn_dep_PLPC; 1.
DR PROSITE; PS50095; PLAT; 1.
DR PROSITE; PS00384; PROKAR_ZN_DEPEND_PLPC; FALSE NEG.
KW Hydrolase; Toxin; Hemolysis; Calcium; Calcium-Binding; Zinc;
KW Signal; 3D-structure.
FT SIGNAL 1 28 BY SIMILARITY.
FT CHAIN 29 398 PHOSPHOLIPASE C.
FT DOMAIN 29 274 PHOSPHOLIPASE C.
FT DOMAIN 275 283 LINKER.
FT DOMAIN 284 398 PLAT.
FT METAL 29 29 ZINC 1.
FT METAL 39 39 ZINC 1.
FT METAL 84 84 ZINC 3.
FT METAL 96 96 ZINC 3.
FT METAL 154 154 ZINC 3.
FT METAL 158 158 ZINC 1 AND 3.
FT METAL 164 164 ZINC 2.
FT METAL 176 176 ZINC 2.
FT METAL 180 180 ZINC 2.
FT METAL 299 299 ZINC 1 (VIA CARBONYL OXYGEN) (BY
FT METAL 300 300 SIMILARITY).
FT METAL 301 301 CALCIUM 3 (BY SIMILARITY).
FT METAL 321 321 CALCIUM 2 (BY SIMILARITY).
FT METAL 322 322 CALCIUM 2 (BY SIMILARITY).
FT METAL 324 324 CALCIUM 2 (VIA CARBONYL OXYGEN) (BY
FT METAL 325 325 SIMILARITY).
FT METAL 326 326 CALCIUM 3 (BY SIMILARITY).
FT METAL 326 326 CALCIUM 2 (BY SIMILARITY).
FT METAL 365 365 CALCIUM 3 (BY SIMILARITY).
SQ SEQUENCE 398 AA; 45598 MW; EF3C95AF603CD5A7 CRC64;
Query Match 15.6%; Score 237; DB 1; Length 398;
Best Local Similarity 26.4%; Pred. No. 1.6e-11;
Matches 79; Conservative 47; Mismatches 113; Indels 60; Gaps 13;
QY 1 MKKYLALAAVVALAAPVQSVVFAQTNNSESPAPILRWSAEDKHNESHILWIVNRAID 60
Db 1 MKRKYLKLLICATITATSLWV-----RTTKVYAM--DGRADGTGTHAMIAQTGV 48
QY 61 I-----MSRNTTIVNPNETALLNWRADLENGIYSADYENPYDINSTYASHFYDPTGT 116
Db 49 ILENDLSSEPEVIRNNLEILKQNMHDLQLGSLTYPDYDKNAYD--LYQHFWDPTDNNF 106
QY 117 IPFAK-----HAKETGAKYFNLAGQAYQNDMQOAFVGLSLHYLDGVNQPMAH 167
Db 107 TKDSKWYLSYISPTDPAESQIRKFSALARYWKRGNYKQATFYLGAMHYFGDAUTPYGAA 166
QY 168 NPTNLSYPMGFHSHYENFVDTIKNNYIVSDSNGYWNKGANPEDWIEGAAVAQKQDPGV 227
Db 167 NVTAVDSP--GHVKETFAEDRKKQYKI-----NTTGSKTNDAFY-SNLTNEDF--- 213

STRAIN-13 / Type A, K211 / Type A, L9 / Type D, NCIB 10663 / Type D, NCIB 10691 / Type B, NCIB 10748 / Type E, and PB6K / Type A; MEDLINE=9609301; PubMed=8522524; Tetsuaki K., Minami J., Matsushita O., Katayama S., Taniguchi Y., Nakamura S., Nishioka M., Okabe A.,
"Phylogenetic analysis of phospholipase C genes from Clostridium perfringens types A to E and Clostridium novyi."; J. Bacteriol. 177:7164-7170(1995).

[8] SEQUENCE FROM N.A., AND USE AS A VACCINE.
STRAIN-13 / Type A, CER 89L43 / Type A, CER 89L1105 / Type A, and CER 89L1216 / Type A;
MEDLINE=96146062; PubMed=8581165; Ginter A., Williamson E.D., Dessy P., Coppe P., Bullifant H., Howells A.M., Titball R.W.;
"Molecular variation between the alpha-toxins from the type strain (NCTC 8237) and clinical isolates of Clostridium perfringens associated with disease in man and animals."; Microbiology 142:191-198(1996).

[9] SEQUENCE FROM N.A.
STRAIN-K21340 / Type A;
MEDLINE=96336195; PubMed=8709860; Kanehama K., Matsushita O., Katayama S.-I., Minami J., Maeda M., Nakamura S., Okabe A.;
"Analysis of the phospholipase C gene of Clostridium perfringens K21340 isolated from Antarctic soil."; Microbiol. Immunol. 40:255-263(1996).

[10] SEQUENCE FROM N.A.
STRAIN-13 / Type A;
MEDLINE=21664373; PubMed=11792842; Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A., Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
"Complete genome sequence of Clostridium perfringens, an anaerobic flesh-eater."; Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).

[11] MUTAGENESIS OF HISTIDINE RESIDUES.
STRAIN-NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
MEDLINE=95173092; PubMed=7868589; Nagahama M., Okagawa Y., Nakayama T., Nishioka E., Sakurai J.;
"Site-directed mutagenesis of histidine residues in Clostridium perfringens alpha-toxin."; J. Bacteriol. 177:1179-1185(1995).

[12] MUTAGENESIS OF ZINC-BINDING, SUBSTRATE-BINDING, AND CATALYTIC RESIDUES.
STRAIN-8-6 / Type A;
MEDLINE=96294749; PubMed=8698464; Guillaud I., Garnier T., Cole S.T.;
"Use of site-directed mutagenesis to probe structure-function relationships of alpha-toxin from Clostridium perfringens."; Infect. Immun. 64:2440-2444(1996).

[13] MUTAGENESIS OF THR-300, AND CONSERVED ASPARTATE AND TYROSINE RESIDUES.
STRAIN-8-6 / Type A;
MEDLINE=98086094; PubMed=9426125; Guillaud I., Alzari P.M., Saliou B., Cole S.T.;
"The carboxy-terminal C2-like domain of the alpha-toxin from Clostridium perfringens mediates calcium-dependent membrane recognition."; Mol. Microbiol. 26:867-876(1997).

[14] MUTAGENESIS OF ASP-297; TYR-303; TYR-335; TYR-359 AND ASP-364.
STRAIN-NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
MEDLINE=20389543; PubMed=10931204; Alape-Giron A., Flores-Diaz M., Guillaud I., Naylor C.E., Titball R.W., Rucavado A., Lomonte B., Basak A.K., Gutierrez J.M., Cole S.T., Thelestam M.;
"Identification of residues critical for toxicity in Clostridium perfringens phospholipase C, the key toxin in gas gangrene."; Eur. J. Biochem. 267:5191-5197(2000).

15] MUTAGENESIS OF ASP-321; ASP-333 AND LYS-358
RC STRAIN=NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
RX MEDLINE=21023335; PubMed=11147832;
RA Walker N., Holley J., Naylor C.E., Flores-Diaz M., Alape-Giron A.,
RA Carter G., Carr F.J., Thelestam M., Moss D.S., Basak A.K.,
RA Miller J., Titball R.W.;
RT "Identification of residues in the carboxy-terminal domain of
RT Clostridium perfringens alpha-toxin (phospholipase C) which are
RT required for its biological activities.";
RL Arch. Biochem. Biophys. 384:24-30(2000).
[16]
RC STRAIN=NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
RX STRAIN=NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
RX MEDLINE=21233265; PubMed=11334886;
RA Jepsen M., Bullifant H.B., Crane D.T., Flores-Diaz M., Alape-Giron A.,
RA Jayasekera P., Lingard B., Moss D.S., Titball R.W.;
RT "Tyrosine 331 and phenylalanine 334 in Clostridium perfringens
RT alpha-toxin are essential for cytotoxic activity.";
RL FEBS Lett. 495:172-177(2001).
[17]
RC STRAIN=NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
RX STRAIN=NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
RX MEDLINE=89248307; PubMed=2497921;
RA Fujii Y., Sakurai J.;
RT "Contraction of the rat isolated aorta caused by Clostridium
RT perfringens alpha toxin (phospholipase C): Evidence for the
RT involvement of arachidonic acid metabolism.";
RL Br. J. Pharmacol. 97:119-124(1989).
[19]
RC STRAIN=NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
RX STRAIN=NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
RX MEDLINE=90255924; PubMed=2111259;
RA Titball R.W., Rubidge T.;
RT "The role of histidine residues in the alpha toxin of Clostridium
RT perfringens.";
RL FEMS Microbiol. Lett. 56:261-265(1990).
[20]
RC STRAIN=NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
RX STRAIN=NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
RX MEDLINE=91209950; PubMed=1902199;
RA Titball R.W., Leslie D.L., Harvey S., Kelly D.C.;
RT "Hemolytic and sphingomyelinase activities of Clostridium perfringens
RT alpha-toxin are dependent on a domain homologous to that of an enzyme
RT from the human arachidonic acid pathway.";
RL Infect. Immun. 59:1872-1874(1991).
[21]
RC STRAIN=NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
RX STRAIN=NCTC 8237 / ATCC 13124 / CN 1491 / Type A;
RX MEDLINE=99386893; PubMed=10456947;
RA Ellmore D.M., Baird R.N., Awad M.M., Boyd R.L., Rood J.I.,
RA Emmins J.J.;
RT "Use of genetically manipulated strains of Clostridium perfringens
RT reveals that both alpha-toxin and theta-toxin are required for
RT vascular leukostasis to occur in experimental gas gangrene.";
Query Match 15.0%; Score 229; DB 1; Length 398;
Best Local Similarity 27.1%; Pred. No. 6.8e-11;
Matches 82; Conservative 44; Mismatches 109; Indels 68; Gaps 15;
QY 1 MKKKVLAAMVALA---APQSVVFAQTNNSESPATLRWSAEKHNESHLMWYN 56
DB 1 MKKKKCALCATLATSLLWAGASTKYVA-----W---DGKIDGTGTHAMIVT 44
QY 57 RAIDMSRNTTIVNPNET-----ALLNWRADLENGIYSADYENPYDNTYASHFVDPPT 112
DB 1 RAIDMSRNTTIVNPNET-----ALLNWRADLENGIYSADYENPYDNTYASHFVDPPT 112

45 QGVSIENDLSKNEPESVRKNLEILKENMHELOGLQSTYPDYDKNAVYD--LYQDHFWDPDT 102
113 GTTYP-----FAKHAKETG-----AKYFNLAGQAYQNMDOQAPFFYLGSLSLHYLGDVNPQ 163
103 DNNFSKNSWLYAISIPDTGESQIRKFSALARYEWQRCNTYKQATFYLGEMHHYFGDIDTP 162
164 MHAANFTNLSPMGFHSKYENFVDITIKNNYIVSDSNGYMNWKGANPEDWIEGAATAAKQD 223
163 YHPANVT--AVDSAGHVKFETFAERKEQYKINTAGCKTN-----EDFY--ADILKNKD 212
224 YPGVNDTTQDW-----FVKAASVQBYA-----DKW--RAEYVTPVTGKRLMEARQVTAG 270
213 P-----NAWSKEYARGFAKTGKSIYYSHASHSHSDDDWDYAAKVT-----LANSQKGTAG 262
271 YIH 273
263 YII 265

RESULT 8
PHLC_CLOHA STANDARD; PRT; 399 AA.
ID PHLC_CLOHA
AC PS9026;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phospholipase C precursor (EC 3.1.4.3) (PLC) (Phosphatidylcholine
DE cholinephosphohydrolase) (Beta toxin).
GN PLC.
OS Clostridium haemolyticum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=84025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=71170;
RA Hauer P.J., Rosenbusch R.F.;
RT "Cloning and molecular characterization of the beta toxin
RT (phospholipase C) gene of Clostridium haemolyticum.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVIEW.
RX MEDLINE=20465460; PubMed=11008117;
RA Jepsen M., Titball R.W.;
RT "Structure and function of clostridial phospholipases C.";
RL Microbes Infect. 2:1277-1284(2000).
CC -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC cell membranes and cause cell rupture. Binds to eukaryotic
CC membranes where it hydrolyzes phosphatidylcholine, sphingomyelin
CC and phosphatidylethanolamine. The diacylglycerol produced can
CC activate both the arachidonic acid pathway, leading to modulation
CC of the inflammatory response cascade and thrombosis, and protein
CC kinase C, leading to activation of eukaryotic phospholipases and
CC further membrane damage (By similarity).
CC -!- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
CC diacylglycerol + choline phosphate.
CC -!- COFACTOR: Binds 3 zinc and 3 calcium ions (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- DOMAIN: The protein is composed of 2 domains; the N-terminal
CC domain contains the phospholipase C active site (PLC), in a cleft
CC which is also occupied by the 3 zinc ions. The C-terminal domain
CC is a putative phospholipid-recognition domain, which shows
CC structural homology with phospholipid-binding C2-like domains from
CC a range of eukaryotic proteins. The ability to bind membrane
CC phospholipids in a Ca(2+) dependent manner and toxicity is
CC conferred by this C-terminal domain, which also contributes to the
CC sphingomyelinase activity (By similarity).
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL ZINC-METALLOPHOSPHOLIPASES C
CC FAMILY.
CC -!- SIMILARITY: Contains 1 PLAT domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

Query Match	Best Local Similarity	Score	DB 1; Length	DB 2; Length
EMBL; AF525415; AM88377.1; -	29	399	PHOSPHOLIPASE C.	29
InterPro; IPR001024; Lipoxigenase_LH2.	29	399	PHOSPHOLIPASE C.	29
InterPro; IPR008976; PLAT LH2.	29	399	PHOSPHOLIPASE C.	29
InterPro; IPR008947; PLC Nuclease.	275	282	LINKER.	275
InterPro; IPR001531; ZN dep PLPC.	283	399	PLAT.	283
Pfam; PF00882; ZN dep PLPC; 1.	28	399	PLAT.	28
PRINTS; PR00479; PRPHLPASEC.	28	399	PLAT.	28
ProDom; PD003946; ZN dep PLPC; 1.	28	399	PLAT.	28
PROSITE; PS50095; PLAT; 1.	28	399	PLAT.	28
PROSITE; PS00384; PROKAR ZN DEPEND PLPC; FALSE NEG.	28	399	PLAT.	28
Hemolysis; Hydrolase; Calcium; Calcium-binding; Zinc; Toxin; Signal.	28	399	PLAT.	28
SIGNAL.	28	399	PLAT.	28
CHAIN	29	399	PHOSPHOLIPASE C.	29
DOMAIN	29	399	PHOSPHOLIPASE C.	29
DOMAIN	275	282	LINKER.	275
DOMAIN	283	399	PLAT.	283
METAL	28	399	PLAT.	28
METAL	28	399	PLAT.	28
METAL	38	38	ZINC 1 (BY SIMILARITY).	38
METAL	83	83	ZINC 1 (BY SIMILARITY).	83
METAL	95	95	ZINC 3 (BY SIMILARITY).	95
METAL	153	153	ZINC 3 (BY SIMILARITY).	153
METAL	157	157	ZINC 3 (BY SIMILARITY).	157
METAL	163	163	ZINC 2 (BY SIMILARITY).	163
METAL	175	175	ZINC 2 (BY SIMILARITY).	175
METAL	179	179	ZINC 2 (BY SIMILARITY).	179
METAL	298	298	CALCIUM 1 (VIA CARBONYL OXYGEN) (BY SIMILARITY).	298
METAL	299	299	CALCIUM 3 (BY SIMILARITY).	299
METAL	300	300	CALCIUM 3 (BY SIMILARITY).	300
METAL	320	320	CALCIUM 2 (BY SIMILARITY).	320
METAL	321	321	CALCIUM 2 (BY SIMILARITY).	321
METAL	323	323	CALCIUM 2 (VIA CARBONYL OXYGEN) (BY SIMILARITY).	323
METAL	324	324	CALCIUM 3 (BY SIMILARITY).	324
METAL	325	325	CALCIUM 2 (BY SIMILARITY).	325
METAL	325	325	CALCIUM 3 (BY SIMILARITY).	325
METAL	363	363	CALCIUM 1 (BY SIMILARITY).	363
SEQUENCE	399	AA; 45942 MW; 97A8FA5537083BEA CRC64;		
Query Match	14.0%;	Score 213.5;	DB 1; Length 399;	
Best Local Similarity	25.9%;	Pred. No. 1.1e-09;		
Matches	7%;	Conservative 40;	Mismatches 108;	Indels 69; Gaps 14;
2	KKKVLALAAVMAALAPVQSVFQAQTNNSESPAPILRWSAEDKHNEGINSHLMTVNRDAID	61		
3	KKKILKFCISAVLS-----FLPSGYSYA-----W-----DGKVDGTGTHALIVTQAVEI	48		
62	MGRTNTIYNP-----NETALLNEWRADLENGTIVSADYENPYDNSTVASHFYDPDGTCTTY	117		
49	LKNDVISTSPSVKENFKILESNLKQLRGSTPYDPRAY--ALYQDFWDPDPTDNN--	104		
118	PPAKHAK-----ETG-----AKYFNLAGVQNMDOQAFFYGLSLHYLGDVNPQMH	165		
105	FTKUSKWYLVAGINETSQSQRKLPALAKDEWKNGYEQATWLLQGLHYFGDFHTPYH	163		
166	AAFTNTLSYPMGFHSHYENFVDTIKNNYIVSDNSGYNWKNGANPEDWIEGAAVAAKQDYP	225		
164	PSNVT--AVDSAGHTKFEYVEGKDSYKLHTA-----GAN-----SVKEFPY	204		
226	GVVNDTTKD-WFVKAASVQEYADKWPAEFTVPTVGKRLMEARQVTAQYIHLWFD	277		
205	TTLQNTLNLDNWT-----EYGRW-----AKKANWYAHATMSHSWKD	243		

RESULT 9

PHLC_CLONO STANDARD; PRT; 398 AA.

Q46150;
28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Phospholipase C precursor (EC 3.1.4.3) (PLC) (Phosphatidylcholine
cholophosphohydrolase) (gamma-toxin).
PLC.
Clostridium novyi.
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
NCBI_TaxID=1542;
[1]
SEQUENCE FROM N.A.
STRAIN=CL49 / Type A;
MEDLINE=96099301; PubMed=8522524;
TsuLeui K., Minami J., Matsuhashita O., Katayama S.-I., Taniguchi Y.,
Nakamura S., Nishioka M., Okabe A.;
"Phylogenetic analysis of phospholipase C genes from Clostridium
perfringens types A to E and Clostridium novyi.";
J. Bacteriol. 177:7164-7170(1995).
[2]
CHARACTERIZATION.
STRAIN=IID 140 / Type A;
MEDLINE=76040159; PubMed=241423;
Taguchi R., Ikezawa H.;
"Phospholipase C from Clostridium novyi type A. I.";
Biochim. Biophys. Acta 409:75-85(1975).
[3]
REVIEW.
MEDLINE=20465460; PubMed=11008117;
Jepson M., Titball R.W.;
"Structure and function of clostridial phospholipases C.";
Microbes Infect. 2:1277-1284(2000).
-!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
cell membranes and cause cell rupture. Binds to eukaryotic
membranes where it hydrolyzes phosphatidylcholine, sphingomyelin
and phosphatidylethanolamine. The diacylglycerol produced can
activate both the arachidonic acid pathway, leading to modulation
of the inflammatory response cascade and thrombosis, and protein
kinase C, leading to activation of eukaryotic phospholipases and
further membrane damage (By similarity). This enzyme is hemolytic
against horse erythrocytes.
-!- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
diacylglycerol + choline phosphate.
-!- COPACITOR: Binds 3 zinc and 3 calcium ions (By similarity).
-!- SUBCELLULAR LOCATION: Secreted (By similarity).
-!- DOMAIN: The protein is composed of 2 domains; the N-terminal
domain contains the phospholipase C active site (PLC), in a cleft
which is also occupied by the 3 zinc ions. The C-terminal domain
is a putative phospholipid-recognition domain, which shows
structural homology with phospholipid-binding C2-like domains from
a range of eukaryotic proteins. The ability to bind membrane
phospholipids in a Ca(2+) dependent manner and toxicity is
conferred by this C-terminal domain, which also contributes to the
sphingomyelinase activity.
-!- SIMILARITY: BELONGS TO THE BACTERIAL ZINC-METALLOPHOSPHOLIPASES C
FAMILY.
-!- SIMILARITY: Contains 1 PLAT domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; D32125; BAA06851.1; -
InterPro; IPR001024; Lipoxigenase_LH2.
InterPro; IPR008976; PLAT_LH2.
InterPro; IPR008947; PLC_Nuclease.
InterPro; IPR001531; Zn_dep_PLPC.


```
Db 314 -----HYFNLTGANYSSQDI-----LYVDH-----HYEVFTD 340
QY 172 LSPYMGFHSKYENFV---DTIKNNY-----IVSDNGYWNWKGANPE--DWIEGAARAAK 221
Db 341 AQLAETQFARIENIYNGDSIHKELSFHAPVGE-----WSGATDCATWLVGVGVAR 394
QY 222 QD--YGVVNDTTKDFVKAAYQSEYADKWRAEVPTVTKRLMEARQV 268
Db 395 YDGSYNTTFTTNDPKVGTCTISQNSLADWTQDYDRV-RQPIEAQLAT 442

RESULT 11
MOKD SCHPO
ID MOKD SCHPO STANDARD; PRT; 2358 AA.
AC Q9Y7I9; O94638;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell wall alpha-1,3-glucan synthase mok13 (EC 2.4.1.183).
GN MOK13 OR SPBC16D10.05.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Katayama S., Dai H., Arellano M., Perez P., Toda T.;
RT "Fission yeast alpha-glucan synthase Moki localizes closely with actin
RT and play a role essential for cell morphogenesis and protein kinase C
RT function.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Guillermo R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard J., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RL "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880 (2002).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {alpha-D-glucosyl-(1,3)} (N) =
CC UDP + {alpha-D-glucosyl-(1,3)} (N+1)
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC
```

```
CC or send an email to license@isb-sib.ch.
CC
DR EMBL; AB018382; BAA76559.1; -
DR EMBL; AL035637; CAB38509.1; -
DR PIR; T39569; T39569.
DR PIR; T43432; T43432.
DR GenBank; SPombe; SPBC16D10.05; -
DR InterPro; IPR006047; Alpha_amy_cat.
DR Pfam; PF00128; alpha-amyase; 1.
KW Cell wall; Transferase; Glycosyltransferase.
FT CONFLICT 120 120 V -> VRRVMLCSLTNKV (IN REF. 1).
SQ SEQUENCE 2358 AA; 269192 MW; A175577C9D8AD731 CRC64;

Query Match 6.7%; Score 102; DB 1; Length 2358;
Best Local Similarity 18.4%; Pred. No. 6.9;
Matches 67; Conservative 43; Mismatches 115; Indels 140; Gaps 15;

QY 3 KKVLAALAAWALAAPVQSVVPAQ-----TNNSESPAPILRWSAEDKHNEGINSHLWIV 55
Db 4 KNILVNLILSIPRLVFTAKYDERESLWNLNQNSATDPLDYWGKWNHQYHPSDDWQV 63
QY 56 NRAIDMSRNTTIVNPNETALINWRADLENGIYSADYENPYDNTSYASHEFYDPTGTT 115
Db 64 -----PFYTVILDKWK-----DGDPRNNEANTIVEYDIY----- 93
QY 116 YIPFAKHAKETGAKYFNLAGQAYONQDMQQAFFYGLSLHYLGD---VNQPMHAANFTNL 172
Db 94 -----ETG---FRNGDIIIGLGLSLDYLEIMGIKIVYIAGTFFPLNQPMGADQYSPL 141
QY 173 SYPM-----GFHSKYENFVDTI----- 189
Db 142 DYTLDHHSHTAGTQWRDTIEIHRGFGYLVLDLTITLIGDLIGFKYLNSTTFFSLFEHE 201
QY 190 ---KNNYIVSDSNGYWNW-----KGANPEDWIE-GAAVAAKQDYPGVVND----- 230
Db 202 AVWKSNIYDP-----WNFTNKYDPCELPRFWGEDGAPVWI--DIVGCVSDSDFOYGDTE 255
QY 231 ---TTKW---FVKAASVQSEYADKWRAEVT-----PVTGRLMEARQV 270
Db 256 AFGTHPDERQLSKFASVQDRLREWRPSVSEKLKHFACMIITAMLDVDGFRIDKATQITVD 315
QY 271 YIHLW 275
Db 316 FLASW 320

RESULT 12
OMPF SALT1 STANDARD; PRT; 363 AA.
ID OMPF SALT1
AC Q56113;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer membrane protein F precursor (Porin ompF) (Outer membrane
DE protein S3).
GN OMPF OR OMP53 OR STY1002 OR T1935.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IMSS-1;
RA Fernandez-Mora M., Calva E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Conneron P.,
RA Cronin A., Davis P., Davies R.M., Dowd N., White N., Farrar J.,
```

RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogsh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TY2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -!- FUNCTION: ompF is a porin that forms passive diffusion pores which
 CC allow small molecular weight hydrophilic materials across the
 CC outer membrane. It is also a receptor for the bacteriophage T2 (By
 CC similarity).
 CC -!- SUBUNIT: Homotrimer (By similarity).
 CC -!- SIMILARITY: Belongs to the Gram-negative porin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X89757; CA61905.1; -;
 DR EMBL; AL627268; CAD05399.1; -;
 DR EMBL; AE016840; AAO69550.1; -;
 DR HSSP; P02931; 1GFN.
 DR InterPro; IPR003229; OMP 2.
 DR InterPro; IPR001702; Porin Gram-ve.
 DR Pfam; PF00267; Gram-ve_porins; 1.
 DR PRINTS; PR00182; ECOLNEIPORIN.
 DR ProDom; PD000808; OMP 2; 1.
 DR PROSITE; PS00576; GRAF NEG PORIN; 1.
 KW Outer membrane; Transmembrane; Porin; Phage recognition; Signal;
 KW Complete proteome.
 FT SIGNAL 1 22 BY SIMILARITY.
 FT CHAIN 23 363 OUTER MEMBRANE PROTEIN F.
 FT TRANSMEM 23 363 BY SIMILARITY.
 FT DOMAIN 29 29 PERIPLASMIC (BY SIMILARITY).
 FT TRANSMEM 30 45 BY SIMILARITY.
 FT DOMAIN 46 56 EXTRACELLULAR (BY SIMILARITY).
 FT TRANSMEM 57 69 BY SIMILARITY.
 FT DOMAIN 70 71 PERIPLASMIC (BY SIMILARITY).
 FT TRANSMEM 72 84 BY SIMILARITY.
 FT DOMAIN 85 99 EXTRACELLULAR (BY SIMILARITY).
 FT TRANSMEM 100 108 PERIPLASMIC (BY SIMILARITY).
 FT DOMAIN 109 109 PERIPLASMIC (BY SIMILARITY).
 FT TRANSMEM 110 117 BY SIMILARITY.
 FT DOMAIN 118 154 EXTRACELLULAR (BY SIMILARITY).
 FT TRANSMEM 155 161 BY SIMILARITY.
 FT DOMAIN 162 169 PERIPLASMIC (BY SIMILARITY).
 FT TRANSMEM 170 181 BY SIMILARITY.
 FT DOMAIN 182 192 EXTRACELLULAR (BY SIMILARITY).
 FT TRANSMEM 193 203 BY SIMILARITY.
 FT DOMAIN 204 204 PERIPLASMIC (BY SIMILARITY).
 FT TRANSMEM 205 217 BY SIMILARITY.
 FT DOMAIN 218 230 EXTRACELLULAR (BY SIMILARITY).
 FT TRANSMEM 231 242 BY SIMILARITY.
 FT DOMAIN 243 243 PERIPLASMIC (BY SIMILARITY).
 FT TRANSMEM 244 256 BY SIMILARITY.
 FT DOMAIN 257 272 EXTRACELLULAR (BY SIMILARITY).
 FT TRANSMEM 273 285 BY SIMILARITY.
 FT DOMAIN 286 287 PERIPLASMIC (BY SIMILARITY).
 FT TRANSMEM 288 301 BY SIMILARITY.

FT DOMAIN 302 312 EXTRACELLULAR (BY SIMILARITY).
 FT TRANSMEM 313 324 BY SIMILARITY.
 FT DOMAIN 325 326 PERIPLASMIC (BY SIMILARITY).
 FT TRANSMEM 327 336 BY SIMILARITY.
 FT DOMAIN 337 353 EXTRACELLULAR (BY SIMILARITY).
 FT TRANSMEM 354 363 BY SIMILARITY.
 FT CONFLICT 307 307 D -> G (IN REF. 1).
 SQ SEQUENCE 363 AA; 40106 MW; P5059B37EA516859 CRC64;

Query Match 6.4%; Score 97.5; DB 1; Length 363;

Best Local Similarity 23.5%; Pred. No. 1.5;

Matches 67; Conservative 28; Mismatches 109; Indels 81; Gaps 17;

QY 1 MKKVLALAAWVAALPQSVVFAQTNNSSPAPILRWSAEDKNEGINSMLTV----- 55

DB 2 MKRKILAAVPAALAAATANA--AEIYNKDG-----NKLDLYGKAVGRHVWTTGDSK 52

QY 56 ---NRAIDIMSRNTTVNPNETALLNEW---RADLENG-----LY 89

DB 53 NADQTYAQIGKGETQINTDLTG-FGQWEYRTKADRAEGEQQNSNLVRLAPAGLKVAEVG 111

QY 90 SADYENPY---YDNSTY---ASHFYDPTGTITYPFAKHAKETGAKYFNLAGQAYQNDM 143

DB 112 SIDYGRNYGIYDVESYTDWAPYSGETWGTGAYTDNYMTSRAGGL-----LTRYNSD- 163

QY 144 QQAEFYLL--GLS--LHYLGDVNPQMAANFTN---LSYPM-----GFHSKYENFVDTI 189

DB 164 ---PFGLVGSLGFIQYQKQ-NQDNHNSINQNGQGVGYTMAYEFGVTVTAAYNSKRT- 218

QY 190 KNNYIVSDSNGYNWKGANPDMTEGAATAAKQDYPGVVNDTKD 234

DB 219 -NQQDRDGN-----GDRAESWAVGAKYDANNVYLAAYVAETRN 256

RESULT 13

OMP_F SALTY

ID OMP_F SALTY STANDARD; PRT; 363 AA.

AC P37432;

DT 01-OCT-1994 (Rel. 30, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Outer membrane protein F precursor (Porin ompF) (Outer membrane

DE protein 1A) (Outer membrane protein 1A) (Outer membrane protein B).

GN OMPF OR STM0999.

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OX NCBI_TaxID=602;

RP SEQUENCE FROM N.A.

RC STRAIN=LT2;

RA Venegas A., Gomez I., Bruce E., Martinez M.;

RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=LT2 / SGSC1412 / ATCC 700720;

EX MEDLINE=21534948; PubMed=11677609;

RA McClelland M., Sanderson K.E.; Spieth J., Clifton S.W., Latreille P.,

RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,

RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,

RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,

RA Waterston R., Wilson R.K.;

RT "Complete genome sequence of Salmonella enterica serovar Typhimurium

RT LT2.";

RL Nature 413:852-856(2001).

CC -!- FUNCTION: OMPF IS A PORIN THAT FORMS PASSIVE DIFFUSION PORES WHICH

CC ALLOW SMALL MOLECULAR WEIGHT HYDROPHILIC MATERIALS ACROSS THE

CC OUTER MEMBRANE. IT IS ALSO RECEPTOR FOR THE BACTERIOPHAGE T2 (By

CC similarity).

CC -!- SUBUNIT: Homotrimer (By similarity).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane

CC (By similarity).

CC -!- SIMILARITY: Belongs to the Gram-negative porin family.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z31594; CAAB3471.1; -;
DR EMBL; AE008743; AAL19933.1; -;
DR PIR; S43159; S43159.
DR HSSP; P02931; IGFN.
DR StyGene; SG10264; OmpF.
DR InterPro; IPR003229; OMP 2.
DR Pfam; PF00267; Gram-ve_porins; 1.
DR PRINTS; PR00182; ECOLNEIPORIN.
DR PRODOM; PD000808; OMP 2; 1.
DR PROSITE; PS00576; GRAM_NEG_PORIN; 1.
KW Outer membrane; Transmembrane; Porin; Phage recognition; Signal;
KW Complete proteome.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 363 OUTER MEMBRANE PROTEIN F.
FT TRANSMEM 23 28 BY SIMILARITY.
FT DOMAIN 29 29 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 30 45 BY SIMILARITY.
FT DOMAIN 46 56 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 57 69 BY SIMILARITY.
FT DOMAIN 70 71 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 72 84 BY SIMILARITY.
FT DOMAIN 85 99 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 100 108 BY SIMILARITY.
FT DOMAIN 109 109 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 110 117 BY SIMILARITY.
FT DOMAIN 118 154 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 155 161 BY SIMILARITY.
FT DOMAIN 162 169 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 170 181 BY SIMILARITY.
FT DOMAIN 182 192 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 193 203 BY SIMILARITY.
FT DOMAIN 204 204 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 205 217 BY SIMILARITY.
FT DOMAIN 218 230 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 231 242 BY SIMILARITY.
FT DOMAIN 243 243 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 244 256 BY SIMILARITY.
FT DOMAIN 257 272 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 273 285 BY SIMILARITY.
FT DOMAIN 286 287 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 288 301 BY SIMILARITY.
FT DOMAIN 302 312 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 313 324 BY SIMILARITY.
FT DOMAIN 325 326 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 327 336 BY SIMILARITY.
FT DOMAIN 337 353 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 354 363 BY SIMILARITY.
FT CONFLICT 232 232 E -> A (IN REF. 1).
FT CONFLICT 232 232 T -> V (IN REF. 1).
SQ SEQUENCE 363 AA; 40048 MW; F6069B34E9516859 CRC64;
Query Match 6.4%; Score 97.5; DB 1; Length 363;
Best Local Similarity 23.5%; Pred. No. 1.5;
Matches 67; Conservative 28; Mismatches 109; Indels 81; Gaps 17;
QY 1 MKKVLALAAWVALAAPVQSVVFAQTNNSESPAPILRWSAEDKHNEGNSHUVIV----- 55
DB 2 MKRKILAAVAPALAAATANA--ABIYKDG-----NKL DLYGKAVGRHWTTTGDGSK 52
QY 56 ---NRAIDMSRNTTIVNPETALLNEW----RADLENG-----IY 89
DB 53 NADQTYAQIGKGETQINTDLTG-FQWEYRTKADRAEGEQNSNLVRLAFAGLKYAEVG 111

QY 90 SADYENPY---YDNSTY---ASHFYDPTGTGTYIPFAKHAKETGAKYFNLAGQAYONQDM 143
DB 112 SIDYGRNYGIYVDVESYTDMAFYSGETWGCAYTDNVTYSRAGGL-----LTYRNSD- 163
QY 144 QQAQFYL--GLS--LHYLGDVNPQPHAAFNFTN---LSYPM-----GFHSKYENFVDTI 189
DB 164 ---FFGLVDGLSFGIYQYQK-NQDNHLSINSQNGDGVGYTMAYEPDGFVGTAAYSNKT- 218
QY 190 KNNYIVSDSNGYWNWKGANPEDWIEGAAVAARQDYPGVVNDTTKD 234
DB 219 -NQQDRDGN-----GDRAESWAVGAKYDANNVYLAAYVAETRN 256
RESULT 14
YEFM_ECOLI
ID YEFM_ECOLI STANDARD; PRT; 468 AA.
AC P75733;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ybfm.
GN YBFM OR B0681.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayaishi K., Honjo A.,
RA Ikenoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000172; AAC73775.1; -;
DR EMBL; D90707; BAA35329.1; ALT_INIT.
DR PIR; H64802; H64802.
DR EcoGene; EG13659; ybfm.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 468 AA; 52780 MW; 954B5A778AE1C2E4 CRC64;
Query Match 6.4%; Score 97; DB 1; Length 468;
Best Local Similarity 22.0%; Pred. No. 2.3;
Matches 78; Conservative 46; Mismatches 136; Indels 94; Gaps 20;
QY 2 KKKVLALA-----AWVALAAPVQSVVFAQTNNSESPAPILRWSAE-----DKHNEG 47
DB 7 KRSLTALAAGVTAMSGFMPEARAEFGIDDTLTGGIYVQWRERDKVTDGDKYKTN 66

```

QY 48 INSHLWIVNR-----AIDI-----MSRNTTIVNPNETALL-----NEWRA 82
D 67 LSHSTWNLNLFQSGYADMFGLDIAFTALEMANGDSSHPNIAFSKSNKAYDEWSG 126
QY 83 DLENGI--YSADYENPYDNTYASHFYDDPTGTT-----YIPFAKHAKETGAKY-FN 132
D 127 D-KSGISLYKAAAFKY--GFWARAGYIQGTQTLAPHSFPGTYQGAEGAGNFDYG 183
QY 133 LAGO---AYQNDQQAFFYGLSHYLGVDVNPQMHAAFNLSYPNGFHSKYENFVDTI 189
D 184 DAGALSFSYMTNFKAPWHELEMBEYQND-----KTKVDYLSHFGAKYD-----F 230
QY 190 KNNYIV-----SDSNGYWN--WKGANPEDMIEGAAVAQDYPGV-----VNDTTKD- 234
D 231 KNNFVLEAFAQGAEGYIDQYFAKASYKFDIAGSPLTTSYQYGTGRKVDKDRSVDNLDYGT 290
QY 235 -WFKAAVSQEVAD--KWRAEVTPV--TGKRLMEARQVTAGY-----IHLWFD 277
D 291 AWLQALTFGYRAADVVDLLEGTWYKADGQQGYFLQRMTPTYASSNGRLDIWMD 344

RESULT 15
OSH2 YEAST
ID OSH2 YEAST STANDARD; PRT; 1283 AA.
AC Q12451, P89891;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oxyterol-binding protein homolog 2.
GN OSH2 OR YDL019C OR D2845.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Urrestarazu L.A., Andre B., Viessers S.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SUBCELLULAR LOCATION.
RX MEDLINE=21301806; PubMed=11408574;
RA Levine T.P., Munro S.;
RT "Dual targeting of Oshlp, a yeast homologue of oxysterol-binding
protein, to both the Golgi and the nucleus-vacuole junction.";
RL Mol. Biol. Cell 12:1633-1644 (2001).
RN [3]
RP GENETIC ANALYSIS.
RX MEDLINE=21135676; PubMed=11238399;
RA Beh C.T., Cool L., Phillips J., Rine J.;
RT "Overlapping functions of the yeast oxysterol-binding protein
homologues.";
RL Genetics 157:1117-1140 (2001).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. When bound to oxysterols it
translocate to the periphery of Golgi membranes.
CC -1- SIMILARITY: Belongs to the OSBP family.
CC -1- SIMILARITY: Contains 2 ANK repeats.
CC -1- SIMILARITY: Contains 1 PH domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z74066; CAA98577.1; -
CC EMBL; Z74067; CAA98578.1; -
CC EMBL; Z48432; CAA86340.1; -
CC PIR; S52500; S52500.
CC GeneOnline; 140261; -
CC SGD; S0002177; OSH2.

```

```

DR GO; GO:0005935; C:bud neck; IDA.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0006594; P:steroid biosynthesis; IGI.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000648; Oxyterol_BP.
DR InterPro; IPR001849; PH.
DR Pfam; PF00203; ank; 2.
DR Pfam; PF01237; Oxyterol_BP; 1.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00248; ANK; 2.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 2.
DR PROSITE; PS01013; OSBP; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Lipid transport; Transport; ANK repeat; Repeat; Golgi stack.
FT REPEAT 106 134
FT REPEAT 206 235
FT DOMAIN 289 386
SQ SEQUENCE 1283 AA; 145795 MW; D521957460E7F7C3 CRC64;

Query Match 6.3%; Score 95.5; DB 1; Length 1283;
Best Local Similarity 25.9%; Pred. No. 11;
Matches 72; Conservative 31; Mismatches 104; Indels 71; Gaps 17;

QY 1 MKKKVLALAAWALAPVQSVVFAQTNNSESPAPILRWSAEDKHNEGINSH----- 51
D 838 VKETKEIASVIGEKTIIVAVTTVQKKEV---LLKEGSLVGEDGIRKRLSKMDKDRP 894
QY 52 ---LWIVNRAI--DIMSRTTIVNPNR-TALLNEWRADLENGIYSADYENPYDNTSYAS 105
D 895 KISLWALVKSMVGKDMTRMTLPVTFNPTSLQORVAEDLE-----YSELLDQAA 943
QY 106 HFYDPDPTGTYI-----PEAKHAKETGAKYFN-LAQQA--YQNDQQAFFYGLSHY 156
D 944 TFDSTLTTLTYAAFTASSYASTTKRV-AKFPNPLLGTEFTEYSRDPQYRFFTEQVSHH- 1001
QY 157 LGDVNQPMHAANFTNLSYPMGFHSKYENFVDTTKNNYIVSDSN-GYWN----- 203
D 1002 -----PPISATWTE-SPRWDFWG--ESFVDTKFGSRFNVKHLGLWHIKLRPNDEKEE 1052
QY 204 ---WKGANPEDMIEGAAVAQDYPG---VNDTTKD 234
D 1053 LYTWK--KPNTVTIGILIGNPQVDNHGENVVNHRTG 1088

Search completed: July 27, 2004, 14:46:31
Job time : 15 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2004, 14:31:12 ; Search time 54 Seconds
(without alignments)
1475.526 Million cell updates/sec

Title: PCT-US03-12556-2
Perfect score: 1522
Sequence: 1 MKKKVLAMVALAAPVQS.....EAQRTAGYIHLWFDYVNR 282

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	511	33.6	289	5	ABB47676 Listeria
2	220	14.5	370	2	AAR52563 Clostridi
3	108.5	7.1	2140	6	ABU01020 S. pneumo
4	108.5	7.1	2140	6	ABU45746 Protein e
5	100	6.6	444	4	ABG08390 Novel hum
6	98	6.4	414	4	AAG0920 Placidia
7	96.5	6.3	836	5	AAG68288 Modified
8	96.5	6.3	836	5	ABU08984 B. latero
9	96.5	6.3	871	5	AAG68285 Brevisbaci
10	96.5	6.3	871	5	ABU08981 B. latero
11	95.5	6.3	1283	6	ABR52701 Protein s
12	94.5	6.2	192	3	AAG52327 Arabidops
13	94.5	6.2	192	3	AAG05009 Arabidops
14	94.5	6.2	305	3	AAG05008 Arabidops
15	94.5	6.2	305	3	AAG52326 Arabidops
16	94.5	6.2	305	5	ABR90956 Herbicida
17	94	6.2	192	2	AAY86159 S. pneumo
18	94	6.2	288	3	AAG10306 Arabidops
19	94	6.2	288	3	AAG24900 Arabidops
20	94	6.2	288	3	AAG50140 Arabidops
21	94	6.2	294	3	AAG24899 Arabidops
22	94	6.2	294	3	AAG50139 Arabidops
23	94	6.2	294	3	AAG10305 Arabidops
24	94	6.2	323	3	AAG24898 Arabidops
25	94	6.2	324	3	AAG50138 Arabidops

26	94	6.2	346	3	AAG10304 Arabidops
27	94	6.2	399	2	AAR53146 C. thermo
28	94	6.2	837	3	AAY70519 Clostridi
29	93.5	6.1	479	5	ABG30886 H. pylori
30	93.5	6.1	787	6	ABU27268 Protein e
31	93	6.1	370	6	ABM72521 Staphyloc
32	93	6.1	664	4	AB48342 S. pneumo
33	93	6.1	2120	3	AAY81710 Streptoco
34	92.5	6.1	304	2	AAR34599 Amaranthu
35	92.5	6.1	304	2	AAR85564 Seed stor
36	92.5	6.1	304	2	AAR83352 Amaranthu
37	92	6.0	1984	4	ABG01338 Novel hum
38	91.5	6.0	238	6	ABM67131 Photorhab
39	90.5	5.9	165	3	AAG05010 Arabidops
40	90.5	5.9	165	3	AAG52328 Arabidops
41	90.5	5.9	296	3	AAG50135 Arabidops
42	90.5	5.9	296	3	AAG05829 Arabidops
43	90.5	5.9	449	5	ABP65417 Bifidobac
44	90.5	5.9	584	6	ABU19765 Protein e
45	90.5	5.9	1320	5	ABG60631 Human pot

ALIGNMENTS

RESULT 1
ABB47676
ID ABB47676 standard; protein; 289 AA.
XX
AC ABB47676;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #380.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR001118.
XX
PR 11-APR-2000; 2000PR-00004629.
XX (INSP) INST PASTEUR.
XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Psihi H, Dehoux P;
XX Dussauguet O, Chetoui F, Nedjati H, Glaser P, Kunst F, Cossart P;
XX Daniels J, Goebel W, Krest J, Kuhn M, Ng E, Vazquez-Roland JA;
XX Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
XX Chakraborty T, Domann B, Hain T, Berche P, Charbit A, Durant L;
XX Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
XX Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
XX Rose M, Voss H;
XX WPI; 2002-010914/01.
DR Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX and prevention of Listeria and related bacterial infections, and related
XX polypeptides.
XX Claim 6; SEQ ID NO 381; 192pp; French.

The present invention relates to the genome sequence of Listeria monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of it are useful for selecting probes and primers for detecting genes in L. monocytogenes and related organisms, and for studying genetic polymorphisms and other genomes. The present sequence is a protein encoded by the genome sequence of the present invention. Proteins

expressed from the genome sequence are useful for raising specific antibodies, identification of L. monocytogenes and related organisms, and for biosynthesis and biodegradation, especially biosynthesis of Vitamin B12. The genome sequence and proteins encoded by it are also useful for selecting compounds that regulate gene expression and cell replication and modulate L. monocytogenes-related diseases. In addition, the genome sequence and proteins encoded by it are useful in pharmaceutical and vaccines compositions for the treatment or prevention of infections by L. monocytogenes and related organisms. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 289 AA;
Query Match 33.6%; Score 511; DB 5; Length 289;
Best Local Similarity 37.1%; Pred. No. 5.7e-39;
Matches 106; Conservative 55; Mismatches 107; Indels 18; Gaps 5;
QY 2 KKKVLAALAMVA-LAAPVQSUV-----FAQT-----NNSESPAP-ILWMSAEDKHNEGIN 49
DB 4 KKVVLGMLIASVLVFPVPTIKANACCDEYLOTPAAPHDIDSKLPHKLSWSADNPTNTDVN 63
QY 50 SHLWTVNRAIDIMSRNTTIVNPNETALLNEWRADLENGIYSADYENPYDNTSYASHPYD 109
DB 64 THYWLFPQAEKILAKDVNHRANLMLNELKKPKDKIAQGIYDADHKNPYDTSFLSHFYN 123
QY 110 PDTGTYTIPFAKHAKETGAKYFNLAGQAYQNQDQQAFFYGLSLHYLGDVNOPMHAANP 169
DB 124 PDRDNTYLPGFANAKITGAKYFNQSVTDYREGKFDTAFFYKGLAIHYTDTISQPMHANF 183
QY 170 TNLSPMPGHSKYENFVDTIKNNYIVSDSNGYNNKGANPEDWISGAAVAQDYPGVVN 229
DB 184 TAISYPPGPHYCHAYENYVDTIKNNYQATQEDMVAKRFCSDDDVLDWLYENAKRAKADIPKLVN 243
QY 230 DTTKDMFVKAASQEVADKWRAEVPVTGKRLMEARQVTAQYIHLW 275
DB 244 ATKXSYLVG-----NSEWKKDVTPTGARLSDSQOTLAGFLFW 283

RESULT 2
ID AARS2563
AC AARS2563 standard; protein; 370 AA.
XX
AC AARS2563;
XX
DT 25-MAR-2003 (revised)
DT 30-MAY-1994 (first entry)
XX
XX Clostridium perfringens alpha-toxin.
DE Alpha-toxin; AT; gas gangrene; vaccine; bacterium; Cpa; conjugate;
XX phospholipase C; sphingomyelin-hydrolysis.
KW
XX Clostridium perfringens.
OS
XX WO9323543-A1.
XX
XX 25-NOV-1993.
XX
XX 20-MAY-1993; 93WO-GB001039.
XX
XX 20-MAY-1992; 92GB-00010717.
PR 23-JUL-1992; 92GB-00015655.
XX
XX (MINA) UK SEC FOR DEFENCE.
XX
XX Titball RW, Williamson ED;
XX
XX WPI; 1993-386575/48.
DR N-PSDB; AAQ62305.
XX
XX New peptides from C-terminus of Clostridium perfringens alpha toxin -

PT useful in protective vaccines against gas gangrene, also derived
PT antibodies for passive immunisation.
XX
XX Disclosure; Page 16-17; 29pp; English.
XX
XX Peptides, or conjugates, have the amino acid sequence of the epitopes of
CC C. perfringens alpha-coxin (AT) from amino acids 261-300 but lack
CC phospholipase C and/or sphingomyelin-hydrolyzing activity. The peptides
CC induce a protective immune response against AT when admin. to human or
CC animals. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 370 AA;

Query Match 14.5%; Score 220; DB 2; Length 370;
Best Local Similarity 26.5%; Pred. No. 1.1e-11;
Matches 68; Conservative 40; Mismatches 95; Indels 54; Gaps 10;
QY 46 EGINSHLWIVNRAIDIMSRNTTIVNPNET-----ALLNEWRADLENGIYSADYENPYDINS 101
DB 6 DGTGTHAMIVTQGVSILENDLSKNPEPSVRKNLEILKENHMLQLGSGTYPDYDKNAYD-- 63
QY 102 TYASHFYDPTGTGYTIP-----FAKHAKETG-----AKYFNLAGQAYQNQDQQAFFYGL 152
DB 64 LYQDFWDPDPTDNNFSDKNWYLAISIPDTGSEQIRKFSALARYEWQRGNYKQATFVLGE 123
QY 153 SLHYLGDVNOPMHAANFNLSPMGHSKYENFVDTIKNNYIVSDSNGYNN----- 203
DB 124 AMHYFGDIDTPYHPANVT--AVDSAGHVKEPTFAEERKEQYKINTAGCKTNEAFYTDILK 181
QY 204 -----WGANPEDWIEGAAVAQDYPGVVNDTTKDMFVKAASQEVAD-KWRAEVPV 256
DB 182 NKDFNMS-----KEYARGFARTGKSIY-----YSHASHMSHSDDDWDAARVTV-- 224
QY 257 TGKRLMEARQVTAQYIHLW 273
DB 225 ----LANSQKGTAGYIY 237

RESULT 3
ID ABU01020
XX ABU01020 standard; protein; 2140 AA.
XX
AC ABU01020;
XX
DT 23-OCT-2003 (revised)
DT 11-FEB-2003 (first entry)
XX
XX S. pneumoniae type 4 strain protein from coding region #590.
DE
XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
XX gene therapy; vaccine.
XX
OS Streptococcus pneumoniae; type 4 strain.
XX
XX WO200277021-A2.
XX
XX 03-OCT-2002.
XX
XX 27-MAR-2002; 2002WO-IB002163.
XX
XX 27-MAR-2001; 2001GB-00007658.
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
XX Masignani V, Tettelin H, Fraser C;
XX
XX WPI; 2003-040579/03.
DR N-PSDB; ABX06302.
XX
XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT useful as medicaments for treating or preventing a disease or infection

PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or ear infection.

XXXXXX

PS Claim 1: SEO ID NO 1180: 56pp: English.

The invention relates to a protein comprising or having at least 50% identity to any of the 2469 amino acid sequences, identified in the specification (available on a computer readable format), or its fragment, expressed from 2469 of 2489 identified DNA coding regions from the *Streptococcus pneumoniae* type 4 strain genomic sequence appearing as A556454. Also included are an antibody which binds one of the proteins, treating a patient by administering the protein, DNA or antibody (in a composition), a kit comprising first and second primers, which are the nucleic acid cited above or fragments between nucleotides 8-100 of a sequence not defined in the specification, for amplifying a target sequence contained within a *Streptococcus* nucleic acid sequence, where the first primer is substantially complementary to the target sequence of the second primer is substantially complementary to the complement of the target sequence, and where the parts of the primers having substantial complementarity define the termini of the target sequence to be amplified, assay comprising contacting a test compound with the protein, and determining whether the test compound binds to the protein and a *Streptococcus pneumoniae* bacterium, where one or more genes encoding the proteins has been rendered inactive. The proteins, nucleic acid molecules, antibody and compositions are useful as medicaments for treating or preventing a disease or infection due to *Streptococcus* bacteria, particularly *S. pneumoniae*, such as pneumonia, sepsis, otitis media or ear infection. They are also useful in developing vaccines, diagnostics and antibiotics. The methods are useful for identifying immunodominant proteins. The present sequence is one of the 2469 proteins expressed by the identified coding regions from the genomic sequence. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to standardise os field)

Db 1 MKKSTVLSITTAAVILAAPNEVVLADTSSSEDALNIDSKEKVAENKEHENIHSAMET 60
Qy 53 -----WIVNRADIMSRNTIYN--PNETALLNEWRADLENGIYADY-----ENP- 96
Db 61 SQDFKEKKTAVIKEXEVSKNPVIDNNTSNEEAKIKEENSNNKSGDYDTSFVNKNTENPK 120
Qy 97 YDNSTYASHFVDPDTGTYTTPFAGHAKETGAKY-----ENLAGOAYO----NQD----- 142
Db 121 KEDKVYIAEFKXESGSKAKELSSLNKTKVLYTYDRIFN--GSAIETTPDNLDKIKQI 178
Qy 143 -----MQQAFYGL--SLHYLGDVNPQMHAANF-----TNLSYPMGF 178
Db 179 EGISSEVERAQVQPMWNNHARKEIGVEAIDYLSINAPF-GKNFDRGWNISNIDTGTDY 237
Qy 179 HSKYENFVDTIK-----NNYIYSD-----SNGYNNWKGANPDWIEGAIVAA 220
Db 238 RHKAMRIDDDAKASMRFKEDLKGTDKNYLSLSDKIPHAFNYNNGKITVEKYDDG----- 292
Qy 221 KQDY-----PGVV--NDTTKD 234
Db 293 -RDVDPHGMHAGILAGNDTEQD 315

RESULT 5

ABG08390
ID ABG08390 standard; protein; 444 AA.

XX ABG08390;

AC ABG08390;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #8381.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

PN 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-P8DB; AAS72577.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

PS Claim 20; SEQ ID NO 38749; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 444 AA;

Query Match 6.6%; Score 100; DB 4; Length 444;

Best Local Similarity 22.3%; Pred. No. 2.1;

Matches 48; Conservative 36; Mismatches 79; Indels 52; Gaps 9;

Qy 2 KKKVLAALAAVMA-----LAAPVQSVVFAQTNNSESPAPILRWSAEDKNEGINSHLWVNR 57

Db 255 KSSILDMGRLISSVENVPKYEGETSSSVSNVSSPIKLM-FLSEVKSEGVKTYLTSGV- 312

Qy 58 AIDIMSRNTIYNPNE---TALLNEWRADLENGIYADYENPYDNDSTYASHFDPDPTGT 114

Db 313 -----TSHSNVLPSEKPTTHVTEETKTETNEDISNSEN-----YHSSHYDTDTFQ 360

Qy 115 TYIPPAKHAKETGAKYFNLAGOAYQNQDMQAFYGLSLHYLGDVN---OPMHAANFT 170

Db 361 RELNKFHAKETA-----GSSTWTFIGDINSDFKQEPKDNSS 397

Qy 171 NLSYPMGFHSHKYENFVDTIKNNIYVSDSNGYNNWK 205

Db 398 SATDP-SFNKXTR---XTKKNSRSPSCETDXATNWK 428

RESULT 6

AAG80920

ID AAG80920 standard; protein; 414 AA.

XX AAG80920;

XX 28-AUG-2001 (first entry)

DT Plastidial delta 9 ACP desaturase partial clone protein sequence.

DE Moss; Physcomitrella patens; lipid metabolism related protein; LMRP;

XX lipid biosynthesis; lipid modification; lipid degradation; cofactor;

XX fatty acid transport; genetic engineering; fatty acid; enzyme; plant;

XX microorganism; polyunsaturated fatty acid; oilseed plant; maize; wheat;

XX biotic stress tolerance; abiotic stress tolerance; rye; oat; triticale;

XX rice; barley; soybean; peanut; cotton; rapeseed; canola; manihot; pepper;

XX sunflower; tagetes; potato; tobacco; eggplant; tomato; Vicia; pea;

XX alfalfa; coffee; cacao; tea; Salix; oil palm; coconut; perennial grass;

XX forage crop.

XX Physcomitrella patens.

OS WO200138484-A2.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-EP011615.

XX 25-NOV-1999; 99WO-EP009108.

XX (BADI) BASF PLANT SCI GMBH.

XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;

PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R;

XX WPI; 2001-367669/38.

XX Nucleic acids encoding lipid metabolism related proteins from

PT Physcomitrella patens useful to produce fine chemicals in modified

PD 26-DEC-2002.
 XX 17-MAY-2001; 2001US-00858525.
 XX 17-MAY-2001; 2001US-00858525.
 XX (BOET/) BOETS A.
 PA (ARNA/) ARNAUT G.
 PA (RIEJ/) RIE J V.
 PA (DAMM/) DAMME N.
 XX Boets A, Arnaut G, Rie JV, Damme N;
 PI WPI; 2003-361872/34.
 DR N-PSDB; ABX94353.
 XX New insecticidal secreted proteins and DNA sequences, useful in making
 PT insecticidal compositions or transgenic plants to protect plants from
 PT damage caused by insects, particularly coleopteran insects.
 XX Claim 4; Page 32-34; 38pp; English.
 XX The invention relates to a new insecticidal secreted protein comprising:
 CC (a) amino acid sequence of the smallest active toxin of the protein
 CC comprising a sequence appearing as ABU08981 (ISP1A) or ABU08982 (ISP2A);
 CC or (b) the amino acid sequence of the protease-digestion fragment of the
 CC protein encoded by the insecticidal secreted protein (isp1a or isp2a) DNA
 CC deposited at the BCM-LMBP under accession number LMBP 4009. Also
 CC included are a DNA sequence encoding the protein cited above, a chimaeric
 CC gene comprising the DNA and a plant-expressible promoter region (e.g. the
 CC long and short forms of the corn zrp2 gene promoter), a plant cell
 CC comprising the chimaeric gene used to produce the protein cited above, a
 CC plant or a seed comprising the chimeric gene of integrated in its cells,
 CC a microorganism transformed to contain the DNA encoding ISP1A or ISP2A, a
 CC process for controlling insects comprising expressing the protein cited
 CC above in cells of a plant, a process for rendering a plant resistant to
 CC coleopteran insects (comprising transforming plant cells with a first
 CC chimaeric gene and a second chimeric gene, and regenerating transformed
 CC plants from the cells which are resistant to insects, where the first and
 CC second chimaeric genes are ISP1A and ISP2A, and a process for controlling
 CC coleopteran insect pests (comprising planting, sowing or growing in a
 CC field plants transformed with a first chimaeric gene and a second
 CC chimaeric gene, where the first and second chimaeric genes are ISP1A and
 CC ISP2A). The insecticidal secreted proteins and DNA sequences are useful
 CC in making insecticidal compositions or transgenic plants to protect
 CC plants (especially crops like corn) from damage caused by insects,
 CC particularly coleopteran insects (e.g. Diabrotica sp., Leptinotarsa sp.
 CC and Anthonomus sp.). The proteins are useful for preventing or minimizing
 CC insect damage to plants in a field. The present sequence represents a
 CC truncated ISP1A protein where the signal peptide has been replaced by Met
 CC -Ala
 XX Sequence 836 AA;
 SQ Query Match 6.3%; Score 96.5; DB 6; Length 836;
 Best Local Similarity 21.0%; Pred. No. 11;
 Matches 64; Conservative 48; Mismatches 140; Indels 53; Gaps 14;
 QY 8 LAAMVALAAPVQSVVFAQT---NNSESAPILRWASAEKHNKNGINSHLWIVNRAIDIM-- 62
 DB 267 VAAFPVNVSLKVLKSKNEDLSHVSQSTNWSY--TNTGVN-----VNAGWSGLGP 319
 QY 63 SRNTIVNPNETALNEWADLENGIYSADYENPY-----YDN-STYASHFYDDPT--- 112
 DB 320 SFGVSVNYQHSETVANWNGSATNDGTHNGASAYLNANRYNNVGTGAIYETKTTTSPF 379
 QY 113 --GTTTIPPAKHAETGAKYFNLAGQYQNMQQAPFFYGLSLHYLGDVNPMPHAAAPT 170
 DB 380 LDGTT-IGTIK-AKENTALTILPDQSYPEKGN-----GIAINTMDDFNSRPIPLNKE 431
 QY 171 NLSYPGFGHSKYENFVDTTKNNYIVSDNG-----YNNWKGANPEDWIEGA-- 218
 DB 432 QLNTYLSNKKPILLETDOVEGKYAIKDTNGNTIAGDMNGITDEISAKTASIIVONGNQ 491

QY 219 -----AAKQDYPGVVNDTTKDMFVKAAVSQSYADKWRARVPTVTGKRLMEARVTVAGIHL 274
 DB 492 SEKRVAADYTN-PEDKTPNLSVKALKLAYPD-----EIEKDGLLFYNDQPIFEASVQS 546
 QY 275 WFDTY 279
 DB 547 YVDEY 551
 RESULT 9
 AAG68285
 ID AAG68285 standard; protein; 871 AA.
 AC AAG68285;
 XX 19-FEB-2002 (first entry)
 DT Brevibacillus laterosporus ISP1A protein SEQ ID NO:2.
 DE Brevibacillus laterosporus; insecticidal secreted protein; ISP1A; ISP2A;
 KW ISP1A-1; ISP2A-1; toxin; bacterial; plant protection.
 XX Brevibacillus laterosporus.
 OS Key Location/Qualifiers
 XX Peptide 1..37
 FT /label= signal
 FT Protein 38..871
 FT /label= ISP1A
 XX WO200187931-A2.
 XX 22-NOV-2001.
 XX 17-MAY-2001; 2001WO-EP005702.
 XX 18-MAY-2000; 2000US-00573872.
 XX (AVET) AVENTIS CROPS SCIENCE NV.
 PA Boets A, Arnaut G, Damme N, Van Rie J;
 PI WPI; 2002-062371/08.
 DR N-PSDB; ABA04056.
 XX New bacterial insecticidal proteins, useful for making insecticidal
 PT composition to protect plants from damage by insects especially
 PT coleopteran insects.
 XX Claim 4; Page 61-65; 86pp; English.
 XX The present sequence represents a Brevibacillus laterosporus insecticidal
 CC secreted protein designated ISP1A. ISP1A, and ISP2A, are insecticidal to
 CC Diabrotica virgifera larvae when ingested by larvae in combination with
 CC the ISP2A amino acid sequence 51 to 457, or ISP1A amino acid sequence 38
 CC to 871, respectively. ISP1A and ISP2A are useful for controlling insects.
 CC The polynucleotide sequences encoding ISP1A and ISP2A can be used in the
 CC production of transgenic plants resistant to coleopteran insects
 XX Sequence 871 AA;
 SQ Query Match 6.3%; Score 96.5; DB 5; Length 871;
 Best Local Similarity 21.0%; Pred. No. 12;
 Matches 64; Conservative 48; Mismatches 140; Indels 53; Gaps 14;
 QY 8 LAAMVALAAPVQSVVFAQT---NNSESAPILRWASAEKHNKNGINSHLWIVNRAIDIM-- 62
 DB 302 VAAFPVNVSLKVLKSKNEDLSHVSQSTNWSY--TNTGVN-----VNAGWSGLGP 354
 QY 63 SRNTIVNPNETALNEWADLENGIYSADYENPY-----YDN-STYASHFYDDPT--- 112
 DB 355 SFGVSVNYQHSETVANWNGSATNDGTHNGASAYLNANRYNNVGTGAIYETKTTTSPF 414

QY 113 --GTTTIPFAKHAKETGAKYFNLAGQAYONQDQQAFFYLGSLHYLGDVNPMAHAAPT 170
DB 415 LDGTT-IGTIK-AKENTTALTILPQSYPEKGN-----GIAINTMDDFNSRPIPLNKE 466
QY 171 NLSYPMGFHSKYENFVDITKNNYIVSDSNG-----YNNWKGANPEDWIEGAIV----- 218
DB 467 QNTYLSNKKPILLETQVEGKYAIKDTNGNITAGDWNGITDEISAKTASIIVDNGNQ 526
QY 219 ----AAKQDYPGVNDTTKDMFVKAASQYADKWRAEVTPVTGKRLMEARQVTAAGYIHL 274
DB 527 SEKRVAAKDYTN-PEDKTPNLSVKEALKLAYPD-----EIEEKDGLLFYNDQPIFEASVQS 581
QY 275 WFDY 279
DB 582 YVDEY 586

RESULT 10
ABU08981
ID ABU08981 standard; protein; 871 AA.
XX AC ABU08981;
XX AC
XX DT 16-JUN-2003 (first entry)
XX DE B. laterosporus insecticidal secreted protein, ISPIA.
XX KW Insecticidal secreted protein; toxin; ISPIA; insecticide; Diabrotica sp;
XX KW Leptinotarsa sp; Anthonomus sp; corn; coleopteran insect; transgenic.
XX OS Brevibacillus laterosporus.
XX FH Key Location/Qualifiers
XX FT Peptide 1..37
XX FT /label= Signal_peptide
XX FT Protein 38..871
XX FT /label= Mature_ISPIA
XX FT /note= "This protein is specifically claimed in claim 2"
XX FT Protein 38..768
XX FT /label= Active_fragment
XX FT /note= "This protein is specifically claimed in claim 3"
XX FN US2002199215-A1.
XX PD 26-DEC-2002.
XX PF 17-MAY-2001; 2001US-00858525.
XX PR 17-MAY-2001; 2001US-00858525.
XX PA (BOET/) BOETS A.
XX PA (ARNA/) ARNAUT G.
XX PA (RIEJ/) RIE J V.
XX PA (DAMN/) DAMME N.
XX PI Boets A, Arnaut G, Rie JV, Damme N;
XX WPI; 2003-361872/34.
XX DR N-PSDB; ABX94348.
XX FT New insecticidal secreted proteins and DNA sequences, useful in making
XX FT insecticidal compositions or transgenic plants to protect plants from
XX FT damage caused by insects, particularly coleopteran insects.
XX PS Claim 1; Page 18-20; 38pp; English.
XX CC The invention relates to a new insecticidal secreted protein comprising:
XX CC (a) amino acid sequence of the smallest active toxin of the protein
XX CC comprising a sequence appearing as ABU08981 (ISPIA) or ABU08982 (ISP2A);
XX CC or (b) the amino acid sequence of the protease-digestion fragment of the
XX CC protein encoded by the insecticidal secreted protein (isPIA or isp2A DNA
XX CC deposited at the BCCM-LMBP under accession number LMBP 4009. Also

CC included are a DNA sequence encoding the protein cited above, a chimaeric
CC gene comprising the DNA and a plant-expressible promoter region (e.g. the
CC long and short forms of the corn zrp2 gene promoter), a plant cell
CC comprising the chimaeric gene used to produce the protein cited above, a
CC plant or a seed comprising the chimeric gene of integrated in its cells,
CC a microorganism transformed to contain the DNA encoding ISPIA or ISP2A, a
CC process for controlling insects comprising expressing the protein cited
CC above in cells of a plant, a process for rendering a plant resistant to
CC coleopteran insects (comprising transforming plant cells with a first
CC chimaeric gene and a second chimeric gene, and regenerating transformed
CC plants from the cells which are resistant to insects, where the first and
CC second chimaeric genes are ISPIA and ISP2A, and a process for controlling
CC coleopteran insect pests (comprising planting, sowing or growing in a
CC field plants transformed with a first chimaeric gene and a second
CC chimaeric gene, where the first and second chimaeric genes are ISPIA and
CC ISP2A). The insecticidal secreted proteins and DNA sequences are useful
CC in making insecticidal compositions or transgenic plants to protect
CC plants (especially crops like corn) from damage caused by insects,
CC particularly coleopteran insects (e.g. Diabrotica sp., Leptinotarsa sp.
CC and Anthonomus sp.). The proteins are useful for preventing or minimising
CC insect damage to plants in a field. The present sequence represents ISPIA
CC protein
XX SQ Sequence 871 AA;

Query Match 6.3%; Score 96.5; DB 6; Length 871;
Best Local Similarity 21.0%; Pred. No. 12;
Matches 64; Conservative 48; Mismatches 140; Indels 53; Gaps 14;
QY 8 LAAMVALAAPVQSVVFAQT---NNSESPAPILRWSAEDKHNGINSHLWVNRADIM-- 62
DB 302 VAAPSVNVSLERKVLKSNEDLSHVSSESQSTNWSY--TNTGVN-----VNAGWSGLGP 354
QY 63 SRNTTIVNPNETALINEWRADLENGIYSADYENPY-----YDN-STYASHFYDPTD--- 112
DB 355 SFGVSVNYQHSSETVANENGSAITNDGTHNGAESAYLANVYNNVGTGAIYETKPTTSFI 414
QY 113 --GTTTIPFAKHAKETGAKYFNLAGQAYONQDQQAFFYLGSLHYLGDVNPMAHAAPT 170
DB 415 LDGTT-IGTIK-AKENTTALTILPQSYPEKGN-----GIAINTMDDFNSRPIPLNKE 466
QY 171 NLSYPMGFHSKYENFVDITKNNYIVSDSNG-----YNNWKGANPEDWIEGAIV----- 218
DB 467 QNTYLSNKKPILLETQVEGKYAIKDTNGNITAGDWNGITDEISAKTASIIVDNGNQ 526
QY 219 ----AAKQDYPGVNDTTKDMFVKAASQYADKWRAEVTPVTGKRLMEARQVTAAGYIHL 274
DB 527 SEKRVAAKDYTN-PEDKTPNLSVKEALKLAYPD-----EIEEKDGLLFYNDQPIFEASVQS 581
QY 275 WFDY 279
DB 582 YVDEY 586

RESULT 11
ABR52701
ID ABR52701 standard; protein; 1283 AA.
XX AC ABR52701;
XX DT 20-JUN-2003 (first entry)
XX DE Protein sequence #SEQ ID 267.
XX KW Multiprotein complex; eukaryote; drug target; diagnosis.
XX OS Saccharomyces cerevisiae.
XX FN EP1258494-A1.
XX PD 20-NOV-2002.
XX PF 20-DEC-2001; 2001EP-00130253.

XX 15-MAY-2001; 2001EP-00111774.
XX (CELL-) CELLZOME AG.
XX Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
XX Marzioch M, Schultz JD, Superti-Furga GD;
XX WPI; 2003-250078/25.
XX N-PSDB; ACC60743.
XX New isolated protein complexes useful for diagnosing a disease or
XX disorder, or as a target for an active agent of a pharmaceutical,
XX preferably a drug target in the treatment or prevention of disease or
XX disorder.
XX Disclosure; SEQ ID NO 267; 17pp + Sequence Listing; English.
XX The invention relates to multiprotein complexes from eukaryotes. Proteins
XX of the invention and DNA sequences encoding them are given in records
XX ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
XX obtainable by using a protein as a bait and isolating the set of proteins
XX which is attached thereto from cells. Such protein complexes may comprise
XX up to 30 distinct proteins. Protein complexes of the invention are useful
XX for diagnosing a disease or disorder, or as a target for an active agent
XX of a pharmaceutical, preferably a drug target in the treatment or
XX prevention of a disease or disorder. Note: The sequence data for this
XX patent is not represented in the printed specification, but is based on
XX sequence information supplied by the European Patent Office. The complete
XX document is available on CD-ROM
XX Sequence 1283 AA;
SQ
Query Match 6.3%; Score 95.5; DB 6; Length 1283;
Best Local Similarity 25.9%; Pred. No. 26;
Matches 72; Conservative 31; Mismatches 104; Indels 71; Gaps 17;
QY 1 MKKVLALAAVLAAPVQSVVPAQTNNSESPATILRWSAEDKHNEGINSH----- 51
DB 838 VKEKTEKTEASSVIGETKIIVAVTTVQKKEEY---LKEGSLYGVEDGKRLSMDKDRP 894
QY 52 ---LWIVNRAI--DIMSRNTTIVNPE-TALLNEWRADLENGIYADYENFYDNTSVAS 105
DB 895 KISLMAVLKSMVKDMTRTLPTFNEPTSLQKVAEDLE-----YSELDDQAA 943
QY 106 HFYDPTGTYYI-----PFAKAKETGAKYFN-LAQQA--YQNDMQQAFFYGLGLSHY 156
DB 944 TFEDSTLRLTYAAFTASSYASTTKRV-AKPFNPLIGETFEYSRDKQYRFTEQVSHH- 1001
QY 157 LGDVNCPMAAFTNLSPMGPHSKYENFVDTIKNNYIVSDGN-GYWN----- 203
DB 1002 -----PPISATWTE-SPRWDFWG--ESFVDTKFNGRSFNVKHLGLWHIKLRPNDEKEE 1052
QY 204 ---WKGANPEDWTEGAARAAKQ-DYPG---VNDTTKD 234
DB 1053 LITWK--KNNVTIGILIGNPQVDNHEVNVVNHHTGD 1088
RESULT 12
AAG52327
ID AAG52327 standard; protein; 192 AA.
XX AAG52327;
AC AAG52327;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 66504.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS

XX PN EP1033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
XX 25-FEB-2000; 2000EP-00301439.
XX PF 99US-0121825P.
XX PR 99US-0123180P.
XX PR 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129645P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 28-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
XX PR 05-MAY-1999; 99US-0132484P.
XX PR 06-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 07-MAY-1999; 99US-0132487P.
XX PR 11-MAY-1999; 99US-0132863P.
XX PR 14-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 14-MAY-1999; 99US-0134370P.
XX PR 18-MAY-1999; 99US-0134768P.
XX PR 19-MAY-1999; 99US-0134941P.
XX PR 20-MAY-1999; 99US-0135124P.
XX PR 21-MAY-1999; 99US-0135353P.
XX PR 24-MAY-1999; 99US-0135629P.
XX PR 25-MAY-1999; 99US-0136021P.
XX PR 27-MAY-1999; 99US-0136392P.
XX PR 28-MAY-1999; 99US-0136782P.
XX PR 01-JUN-1999; 99US-0137222P.
XX PR 03-JUN-1999; 99US-0137528P.
XX PR 04-JUN-1999; 99US-0137502P.
XX PR 07-JUN-1999; 99US-0137724P.
XX PR 08-JUN-1999; 99US-0138094P.
XX PR 10-JUN-1999; 99US-0138540P.
XX PR 10-JUN-1999; 99US-0138847P.
XX PR 14-JUN-1999; 99US-0139119P.
XX PR 16-JUN-1999; 99US-0139452P.
XX PR 16-JUN-1999; 99US-0139453P.
XX PR 17-JUN-1999; 99US-0139492P.
XX PR 18-JUN-1999; 99US-0139454P.
XX PR 18-JUN-1999; 99US-0139455P.
XX PR 18-JUN-1999; 99US-0139456P.
XX PR 18-JUN-1999; 99US-0139457P.
XX PR 18-JUN-1999; 99US-0139458P.
XX PR 18-JUN-1999; 99US-0139459P.
XX PR 18-JUN-1999; 99US-0139460P.
XX PR 18-JUN-1999; 99US-0139461P.
XX PR 18-JUN-1999; 99US-0139462P.
XX PR 18-JUN-1999; 99US-0139463P.
XX PR 18-JUN-1999; 99US-0139750P.
XX PR 18-JUN-1999; 99US-0139763P.
XX PR 21-JUN-1999; 99US-0139817P.
XX PR 22-JUN-1999; 99US-0139899P.
XX PR 23-JUN-1999; 99US-0140353P.
XX PR 23-JUN-1999; 99US-0140354P.
XX PR 24-JUN-1999; 99US-0140695P.
XX PR 28-JUN-1999; 99US-0140823P.
XX PR 29-JUN-1999; 99US-0140991P.
XX PR 30-JUN-1999; 99US-0141287P.

OS Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130444P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132040P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140655P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.


```
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0143977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145114P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147418P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150565P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.

PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159323P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 6.2%; Score 94.5; DB 3; Length 305;
Best Local Similarity 21.3%; Pred. No. 3.9;
Matches 35; Conservative 29; Mismatches 49; Indels 51; Gaps 9;

Qy 135 GQAYQNQDMQQAFFYLGSLHLYLGDVNPQMHAANFTN-----LSYPMGFH---- 179
Db 133 GTSRRYNNMTEALLFLS---HFMDIHQPMH-VGFTSDEGGNTIDLWYKHKSNLHHWD 188
Qy 180 -----SKYENFVDTIKNNYIVSDSNGW-----NWKGAN-----PEDWI-EGAAVA 219
Db 189 REIILTALKENYDKNLDLQEDLEKNTINGLWHDDLSLWTECNDLIACPHKYASESIKLA 248
Qy 220 AKQDYPGVVNDTTKDMFVKAASVOEYADKRAEVPVTVTKGLME 263
Db 249 CKWGYKGVKSGET-----LSEYFN-----TFLPIVMKRIVQ 280

RESULT 15
AAG52326
ID AAG52326 standard; protein; 305 AA.
XX AAG52326;
AC AAG52326;
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 66503.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW
```

KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131448P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149175P.
PR 20-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150586P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151308P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.

```
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 6.2%; Score 94.5; DB 3; Length 305;
Best Local Similarity 21.3%; Pred. No. 3.9;
Matches 35; Conservative 29; Mismatches 49; Indels 51; Gaps 9;

QY 135 GOAYQNDMQCAFFYGLSLHYLDVNPQMHAAFTN-----LSYPMGFH---- 179
Db 133 GTSDRRYNTALLFLS---HPMGDIHQPMH-VGFTSDEGGNTIDLRWYKXSNLHHYWD 188

QY 180 -----SKYENFVDTIKNYIYVDSNGYW-----NWKGAN-----PEDWI-EGAAVA 219
Db 189 REIILTALKENYDKNLDLLQEDLEKNIITNGLMHDLSWTECNLDLIACPHKYASESIKLA 248

QY 220 AKQDYPGVNDTKDWFVKAQVQYADKWRABVTPTVGKRLME 263
Db 249 CKWGYKGVKSGET-----LSEEFN-----TRLPIVMKRIVQ 280
```

Search completed: July 27, 2004, 14:46:07
Job time : 56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2004, 14:45:12 ; Search time 18 Seconds
(without alignments)
808.807 Million cell updates/sec

Title: PCT-US03-12556-2
Perfect score: 1522
Sequence: 1 MKKKVLAMVALAPVQS.....EAQRTAGYIHLNFDYVNR 282

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgm2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgm2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgm2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgm2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgm2_6/ptodata/2/iaa/PCUTS_COMB.pep:*
6: /cgm2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	220	14.5	370	2	US-08-341-538A-2
2	220	14.5	370	2	US-08-725-518-2
3	96.5	6.3	276	4	US-09-510-322A-11
4	95	6.2	274	3	US-08-491-954-4
5	94	6.2	399	2	US-08-282-197C-54
6	94	6.2	685	4	US-09-252-991A-32033
7	94	6.2	837	4	US-09-390-234-14
8	94	6.2	837	4	US-09-603-311-14
9	93	6.1	275	4	US-09-510-322A-10
10	91.5	6.0	372	4	US-09-092-315-13
11	91.5	6.0	464	4	US-09-092-315-1
12	91.5	6.0	464	4	US-09-733-524A-1
13	91.5	6.0	478	4	US-09-092-315-7
14	91.5	6.0	478	4	US-09-733-524A-7
15	91.5	6.0	503	4	US-09-252-991A-17709
16	90	5.9	575	4	US-09-134-001C-3906
17	90	5.9	727	4	US-09-134-001C-4067
18	89.5	5.9	636	3	US-08-737-752A-2
19	89	5.8	218	4	US-09-615-192A-269
20	89	5.8	474	4	US-09-489-039A-13140
21	89	5.8	987	4	US-09-543-681A-7785
22	88.5	5.8	378	4	US-09-489-039A-13960
23	88.5	5.8	423	4	US-09-328-352-5224
24	87.5	5.7	493	4	US-09-543-681A-5068
25	87.5	5.7	515	4	US-09-107-532A-5603
26	87	5.7	325	4	US-09-543-681A-7453
27	86.5	5.7	291	4	US-09-495-406-5

28	86.5	5.7	464	3	US-09-426-072-2	Sequence 2, Appli
29	86	5.7	296	4	US-09-510-322A-2	Sequence 2, Appli
30	85.5	5.6	291	4	US-09-495-406-9	Sequence 9, Appli
31	85	5.6	890	2	US-08-483-101-14	Sequence 14, Appl
32	84.5	5.6	713	4	US-09-489-039A-10668	Sequence 10668, A
33	83.5	5.5	268	4	US-09-134-000C-3519	Sequence 3519, Ap
34	83.5	5.5	418	4	US-09-958-548-1	Sequence 1, Appli
35	83.5	5.5	731	4	US-09-107-532A-6999	Sequence 6999, Ap
36	83.5	5.5	901	4	US-09-134-001C-5389	Sequence 5389, Ap
37	83	5.5	511	1	US-08-480-604A-20	Sequence 20, Appl
38	83	5.5	511	2	US-08-405-496A-20	Sequence 20, Appl
39	83	5.5	511	3	US-08-915-136-20	Sequence 20, Appl
40	83	5.5	511	4	US-08-957-310-20	Sequence 20, Appl
41	83	5.5	511	4	US-10-011-366-20	Sequence 20, Appl
42	83	5.5	511	4	US-09-084-517-20	Sequence 20, Appl
43	83	5.5	608	1	US-08-480-604A-21	Sequence 21, Appl
44	83	5.5	608	2	US-08-405-496A-21	Sequence 21, Appl
45	83	5.5	608	3	US-08-915-136-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-341-538A-2
; Sequence 2, Application US/08341538A
; Patent No. 5817317
; GENERAL INFORMATION:
; APPLICANT: TITBALL DR, RICHARD W
; APPLICANT: WILLIAMSON DR, ETHEL D
; NUMBER OF INVENTION: CLOSTRIDIUM PERFRINGENS VACCINES
; TITLE OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE
; STREET: 8th FLOOR, 1100 NORTH GLEBBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: UNITED STATES
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/341.538A
; FILING DATE: 28-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9210717.6
; FILING DATE: 20-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9215655.3
; FILING DATE: 23-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO GB93/01039
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAWFORD, ARTHUR R
; REGISTRATION NUMBER: 25327
; REFERENCE/DOCKET NUMBER: 1498-58
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 370 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-341-538A-2

Query Match 14.5%; Score 220; DB 2; Length 370;
Best Local Similarity 26.5%; Pred. No. 6.8e-14;
Matches 68; Conservative 40; Mismatches 95; Indels 54; Gaps 10;

```
QY 46 EGINSHLWVNRADIMSRNTTIVNPNET-----ALLNEWRADLENGIYSADYENPYDNS 101
D 6 DGTGTHAMIVTQGVSIENDLSKNEPESVRKNLEILKENMHELQLGSTYPDYDKNAYD-- 63
QY 102 TYASHFYDDPTGTTTYP-----FAKHAKETG-----AKYFNLAGOAYQNQDMQOAFYFYLGL 152
D 64 LYQDHFWDPTDNNFNSKDNSWYLAYSIPDTGESQIRKFSALARYEMQORGNKQATFYLGE 123
QY 153 SLHYLGDVNOPMHAANFTNLSYPMGPHSKYENFVDVTIKNNYIVSDSNGYN----- 203
D 124 AMHYFGDIDTYPHPANVT--AVDSAGHVKFTFAERKEQYKINTAGCKTNEAFTYDILK 181
QY 204 -----WKGANPEWIEGAAVAADQDYPGVVNDTTKOWFVKAASQBYAD-KWRAEYTPV 256
D 182 NKDFNAMS-----KEYARGPAKTGKSIY-----YSHASHSHSDDWDYAAKVT-- 224
QY 257 TGKRLMEARVAGYIH 273
D 225 ----LANSQKGTAGYIY 237

RESULT 2
US-08-725-518-2
; Sequence 2, Application US/08725518
; Patent No. 5851827
; GENERAL INFORMATION:
; APPLICANT: TITBALL DR, RICHARD W
; APPLICANT: WILLIAMSON DR, ETHEL D
; TITLE OF INVENTION: CLOSTRIDIUM PERFRINGENS VACCINES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE
; STREET: 8th FLOOR, 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: UNITED STATES
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/725,518
; FILING DATE: 4-OCT-96
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/341538
; FILING DATE: 28-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9210717.6
; FILING DATE: 20-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9215655.3
; FILING DATE: 23-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO GB93/01039
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAWFORD, ARTHUR R
; REGISTRATION NUMBER: 25327
; REFERENCE/DOCKET NUMBER: 124-536
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 370 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-725-518-2

Query Match 14.5%; Score 220; DB 2; Length 370;
Best Local Similarity 26.5%; Pred. No. 6.8e-14;
```

```
Matches 69; Conservative 40; Mismatches 95; Indels 54; Gaps 10;
QY 46 EGINSHLWVNRADIMSRNTTIVNPNET-----ALLNEWRADLENGIYSADYENPYDNS 101
D 6 DGTGTHAMIVTQGVSIENDLSKNEPESVRKNLEILKENMHELQLGSTYPDYDKNAYD-- 63
QY 102 TYASHFYDDPTGTTTYP-----FAKHAKETG-----AKYFNLAGOAYQNQDMQOAFYFYLGL 152
D 64 LYQDHFWDPTDNNFNSKDNSWYLAYSIPDTGESQIRKFSALARYEMQORGNKQATFYLGE 123
QY 153 SLHYLGDVNOPMHAANFTNLSYPMGPHSKYENFVDVTIKNNYIVSDSNGYN----- 203
D 124 AMHYFGDIDTYPHPANVT--AVDSAGHVKFTFAERKEQYKINTAGCKTNEAFTYDILK 181
QY 204 -----WKGANPEWIEGAAVAADQDYPGVVNDTTKOWFVKAASQBYAD-KWRAEYTPV 256
D 182 NKDFNAMS-----KEYARGPAKTGKSIY-----YSHASHSHSDDWDYAAKVT-- 224
QY 257 TGKRLMEARVAGYIH 273
D 225 ----LANSQKGTAGYIY 237

RESULT 3
US-09-510-322A-11
; Sequence 11, Application US/09510322A
; Patent No. 6391557
; GENERAL INFORMATION:
; APPLICANT: Fox Chase Cancer Center
; TITLE OF INVENTION: Nucleic Acid Encoding a Mismatch
; TITLE OF INVENTION: Endonuclease and Methods of Use Thereof
; FILE REFERENCE: 9503dna
; CURRENT APPLICATION NUMBER: US/09/510,322A
; CURRENT FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Plant
; US-09-510-322A-11

Query Match 6.3%; Score 96.5; DB 4; Length 276;
Best Local Similarity 21.4%; Pred. No. 0.14;
Matches 46; Conservative 31; Mismatches 67; Indels 71; Gaps 12;
QY 103 YASHFY---DPDTGTTTYPFAKHAKETGAKYFNLAGOAYQN-----QDM 143
D 54 WTSPLYIDTDPQACSYEYSDCHDQGLKDMCDVG-AIQNFTSOLQHYGEGTSDRRYNN 112
QY 144 QOAFYFYLGLSLHYLGDVNOPMHAANFTN-----LSYPMGFH----- 179
D 113 TEALLFLS---HPMGDIHQPMH-VTFTSDEGNTIDLRWYKHKSNLHHVWDREIILTALK 168
QY 180 SKYENFVDVTIKNNYIVSDSNGYN-----NWKGAN-----PEDWI-EGAAVAAKQDYPGV 228
D 169 ENYDKNLDLLQEDLEKNITNGLWHDDLSSWTECDNLACPHKAYASESIKLACKMKGKGVK 228
QY 229 NDTTKDWFVKAASQBYADKWRAEYTPVYTGRLME 263
D 229 SGET-----LSBEYFN-----TRLPIVMKRIVQ 251

RESULT 4
US-08-491-954-4
; Sequence 4, Application US/08491954
; Patent No. 6096321
; GENERAL INFORMATION:
; APPLICANT: Girardeau, Jean-Pierre
; APPLICANT: Martin, Christine
; APPLICANT: Mechin, Marie-Claire
; APPLICANT: Der Vartanian, Maurice
; APPLICANT: Bousquet, Francois
```

RESULT 5
US-08-282-197C-54

; PATENT NO. 8331793
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136


```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      July 27, 2004, 14:46:13 ; Search time 47 Seconds
              (without alignments)
              1878.925 Million cell updates/sec

Title:      PCT-US03-12556-2
Perfect score: 1522
Sequence:    1 MKKKVLALAAWALAAPVQS.....EAQRVTAGYIHLWFDITYVNR 282

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:      1288442 seqs, 313154207 residues

Total number of hits satisfying chosen parameters:      1288442

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :      Published Applications AA:*

```

```

Database : Published Applications AA.*
1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1522	100.0	282	15	US-10-421-654-2		Sequence 2, Appl1
2	1257.5	82.6	397	15	US-10-369-493-16747		Sequence 16747, A
3	1232	80.9	283	15	US-10-421-654-4		Sequence 4, Appl1
4	1206.5	79.3	280	15	US-10-421-654-6		Sequence 6, Appl1
5	522	34.3	473	15	US-10-421-654-82		Sequence 82, Appl1
6	515	33.8	473	15	US-10-421-654-90		Sequence 90, Appl1
7	509	33.4	473	15	US-10-421-654-98		Sequence 98, Appl1
8	501	32.9	473	15	US-10-421-654-32		Sequence 32, Appl1
9	120	7.9	251	15	US-10-421-654-106		Sequence 106, App
10	108.5	7.1	2140	12	US-10-282-1228-73670		Sequence 73670, A
11	102.5	6.7	272	12	US-10-424-599-239481		Sequence 239481, A
12	96.5	6.3	836	9	US-09-858-525A-10		Sequence 10, Appl1
13	96.5	6.3	871	9	US-09-858-525A-2		Sequence 2, Appl1
14	96	6.3	293	12	US-10-424-599-219446		Sequence 219446, A
15	93.5	6.1	479	9	US-09-733-524-17		Sequence 17, Appl1

16	93.5	6.1	787	12	US-10-282-1222A-55192	Sequence 55192, A
17	93	6.1	664	13	US-10-067-385-6	Sequence 6, Appli
18	93	6.1	2119	10	US-09-769-744A-28	Sequence 28, Appl
19	92.5	6.1	134	12	US-10-424-599-264682	Sequence 264682,
20	91.5	6.0	372	13	US-10-120-319-13	Sequence 13, Appl
21	91.5	6.0	372	14	US-10-189-977-13	Sequence 13, Appl
22	91.5	6.0	464	13	US-10-120-319-1	Sequence 1, Appli
23	91.5	6.0	464	14	US-10-189-977-1	Sequence 1, Appli
24	91.5	6.0	464	14	US-10-392-098-1	Sequence 1, Appli
25	91.5	6.0	478	13	US-10-120-319-7	Sequence 7, Appli
26	91.5	6.0	478	14	US-10-189-977-7	Sequence 7, Appli
27	91.5	6.0	478	14	US-10-392-098-7	Sequence 7, Appli
28	91.5	6.0	802	15	US-10-369-493-13784	Sequence 13784, A
29	90.5	5.9	584	12	US-10-282-1222A-47689	Sequence 47689, A
30	90.5	5.9	934	15	US-10-369-493-4171	Sequence 4171, Ap
31	90	5.9	270	12	US-10-424-599-239478	Sequence 239478,
32	90	5.9	328	16	US-10-437-963-120169	Sequence 120169,
33	90	5.9	721	12	US-10-282-1222A-70739	Sequence 70739, A
34	90	5.9	1365	16	US-10-383-930-39	Sequence 39, Appl
35	89.5	5.9	636	9	US-09-843-007-2	Sequence 2, Appli
36	89.5	5.9	636	16	US-10-705-195-5	Sequence 5, Appli
37	89	5.8	218	14	US-10-174-693-269	Sequence 269, App
38	89	5.8	866	12	US-10-282-1222A-43305	Sequence 43305, A
39	88.5	5.8	501	9	US-09-733-524-1	Sequence 1, Appli
40	88.5	5.8	517	15	US-10-381-596A-4	Sequence 4, Appli
41	88.5	5.8	887	16	US-10-437-963-191038	Sequence 191038,
42	88.5	5.8	894	12	US-10-282-1222A-77188	Sequence 77188, A
43	88.5	5.8	1786	16	US-10-637-544-20	Sequence 20, Appl
44	88	5.8	1659	16	US-10-474-776-249	Sequence 249, App
45	87.5	5.7	195	12	US-10-424-599-211603	Sequence 211603,

ALIGNMENTS

```

RESULT 1
US-10-421-654-2
; Sequence 2, Application US/10421654
; Publication No. US2004005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421.654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-10-421-654-2

Query Match 100.0%; Score 1522; DB 15; Length 282;
Best Local Similarity 100.0%; Pred. No. 1.9e-139;
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 MKKKVLAALAAWVLAAPQSVVFPAQTNNSSSPAPILKWSAEDKNEGINSHLWVNR 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MKKKVLAALAAWVLAAPQSVVFPAQTNNSSSPAPILKWSAEDKNEGINSHLWVNR 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 61 IMSRNTTIVNPNETALLNEWRADILENGIYSADYENPYDNTSYASHFYDPDGTGTYIPFA 120

```

```
Db 1 IMSRNTTIVNPTALLNEWADLENGIYSADYENPYDNTSYASHFYDPDGTGTYIPPA 120
Qy 121 KHAKETGAKYFNLAGQAYQNDQMOQAPFYGLSLHYLGDVNPQPMHAANFTNLSYPMGFHS 180
Db 121 KHAKETGAKYFNLAGQAYQNDQMOQAPFYGLSLHYLGDVNPQPMHAANFTNLSYPMGFHS 180
Qy 181 KYENFVDITKKNYIVSDSGYNNKMGANPEWIEGAAVAAKQDYPGVVNDTTKDFVCAA 240
Db 181 KYENFVDITKKNYIVSDSGYNNKMGANPEWIEGAAVAAKQDYPGVVNDTTKDFVCAA 240
Qy 241 VQSEYADKWRARVPTVTKRLMEARVVTAGYIHLWFDYTVNR 282
Db 241 VQSEYADKWRARVPTVTKRLMEARVVTAGYIHLWFDYTVNR 282

RESULT 2
US-10-369-493-16747
; Sequence 16747, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16747
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(397)
; OTHER INFORMATION: unsure at all xaa locations
US-10-369-493-16747

Query Match 82.6%; Score 1257.5; DB 15; Length 397;
Best Local Similarity 80.2%; Pred. No. 1.6e-113;
Matches 227; Conservative 23; Mismatches 32; Indels 1; Gaps 1;

Qy 1 MKKKVLALAAVMALAAPVQSVVFAQTNNSESAPIL-RWSAEDKHNEGINSHLWIVNRAI 59
Db 1 MKKKVLALAAITLVAPLQSVAFAHENDGQRFPGVIPRWSAEDKHKEGVNSHLWIVNRAI 60

Qy 60 DIMSRNTTIVNPTALLNEWADLENGIYSADYENPYDNTSYASHFYDPDGTGTYIPF 119
Db 61 DIMSRNTTIVKQDRVALLNEWRTLENGIYAADYENPYDNTSYASHFYDPDNGKTYIPY 120

Qy 120 AKHAKETGAKYFNLAGQAYQNDQMOQAPFYGLSLHYLGDVNPQPMHAANFTNLSYPMGFH 179
Db 121 AKQAKETGAKYFKLAGEYKNDKQAPFYGLSLHYLGDVNPQPMHAANFTNLSYPMQGFH 180

Qy 180 SKYENFVDITKKNYIVSDSGYNNKMGANPEWIEGAAVAAKQDYPGVVNDTTKDFVCAA 239
Db 181 SKYENFVDITKKNYIVSDSGYNNKMGANPEWIEGAAVAAKQDYPGVVNDTTKDFVCAA 240

Qy 240 AVSQEYADKWRARVPTVTKRLMEARVVTAGYIHLWFDYTVNR 282
Db 241 AVSQEYADKWRARVPTVTKRLMEARVVTAGYIHLWFDYTVNR 283

RESULT 3
US-10-421-654-4
; Sequence 4, Application US/10421654
; Publication No. US20040005604A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(24)
US-10-421-654-4

Query Match 80.9%; Score 1232; DB 15; Length 283;
Best Local Similarity 76.5%; Pred. No. 3e-111;
Matches 215; Conservative 35; Mismatches 31; Indels 0; Gaps 0;

Qy 1 MKKKVLALAAVMALAAPVQSVVFAQTNNSESAPIL-RWSAEDKHNEGINSHLWIVNRAID 60
Db 1 MKRKILATASVIALTAPIQSVAFAHENGHPPTIALKWSAESIHNEGVSLSHLWIVNRAID 60

Qy 61 IMSRNTTIVNPTALLNEWADLENGIYSADYENPYDNTSYASHFYDPDGTGTYIPPA 120
Db 61 IMSQNTTIVVQNKTALLNEWRTLEKGIYSADYENPYDNTSYASHFYDPDNGKTYIPPA 120

Qy 121 KHAKETGAKYFNLAGQAYQNDQMOQAPFYGLSLHYLGDVNPQPMHAANFTNLSYPMGFHS 180
Db 121 KQAKQTGAKYFKLAGEAYQNDKNAFPYGLSLHYLGDVNPQPMHAANFTNLSHPFGFHS 180

Qy 181 KYENFVDITKKNYIVSDSGYNNKMGANPEWIEGAAVAAKQDYPGVVNDTTKDFVCAA 240
Db 181 KYENFVDITKKNYIVSDSGYNNKMGANPEWIEGAAVAAKADPSPISVNDKTKNWFCAA 240

Qy 241 VQSEYADKWRARVPTVTKRLMEARVVTAGYIHLWFDYTVN 281
Db 241 VQSDSADKWRARVPTVTKRLMEARVVTAGYIHLWFDYTVN 281

RESULT 4
US-10-421-654-6
; Sequence 6, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
```



```
; OTHER INFORMATION: Obtained from an environmental sample
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-10-421-654-6

Query Match      79.3%; Score 1206.5; DB 15; Length 280;
Best Local Similarity 76.2%; Pred. No. 8.9e-109;
Matches 215; Conservative 34; Mismatches 28; Indels 5; Gaps 2;

Qy 1 MKKKVLALAAAMVALAAPVQS--VVFQATNNSESPAPILRWSAEDKNEGIN 59
Db 1 MKKKLALATASVIALTAIPQSAFAH-----ESDPIALRWSAESVNEGSSHLWVNR 56

Qy 60 DIMSRNTTIVNPNETALLNWRADLENGIYSADYENPYDNTSYASHFYDPDGTGTYIPF 119
Db 57 DIMSQNTTIVKQNETALLNWRNTLEEGIYSADYKPNPYDNTSFASHFYDPDSEKTYIPF 116

Qy 120 AKHAKETGAKYNLAGAYONQDMQAQFFYGLSLHLGLDVNQPMAAANFTNLSYPMGFH 179
Db 117 AKQAKQTGAKYFKLAGEAYQNDLKNAPFYGLSLHLGLDVNQPMAAANFTNISHPPGFH 176

Qy 180 SKYENFVDITKKNYIVSDNSGYNWKNKANPDWIEGAAVAADYDYPGVVNDTTKQWFKVA 239
Db 177 SKYENFVDITKKNYIVSDNSGYNWKNKANPDWIEGAAVAADYDYPGVVNDTTKQWFKVA 236

Qy 240 AVSQEYADKWAERAEVTPVTGKRLMEARVTAQYIHLWFDITYN 281
Db 237 AVSQDSADKWAERAEVTPVTGKRLMEARVTAQYIHLWFDITYN 278

RESULT 5
US-10-421-654-82
; Sequence 82, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; NAME/KEY: SIGNAL
; LOCATION: (1)...(25)
US-10-421-654-82

Query Match      34.3%; Score 522; DB 15; Length 473;
Best Local Similarity 38.7%; Pred. No. 6.7e-42;
Matches 121; Conservative 43; Mismatches 105; Indels 44; Gaps 10;

Qy 1 MKKKVLALAAAMVALAAPVQS--VVFQATNNSESPAPILRWSAEDKNEGIN 49
Db 1 MKKKLCTALVTAISSGVVITPTTEAQACGIGEVKQENQEHK--RVKWSAEHPHHPNES 58

Qy 50 SHLWVNRADIMSRN--TTIVNPNETALLN--EWRADLENGIYSADYENPYDNTST--- 102
Db 59 THLWTRARNAIQIMSRNQKTVQENELQFLNTPYKELFERGLYDADYLDDEFNDGDTG 118

Qy 103 -----YASHFYDPDGTGTYIPPAKHAKE-----TGAKYFNLAGQAYQNDMQQAFFY 149
Db 119 IDGLIKGGWKSHFYDPDTRKNY----KGEETPTALSQGDQKFKLAGDYFKKEDQQAFFY 174

Qy 150 LGLSLHLYLGDVNPQMAAANFTNL--SYPMGFHSKYENFVDITKKNYIVSDNSGYNWKNKAN 208
Db 175 LGVATHYFTDATQPMHAANFTAVDTSAKFKHSAFENYVTTIQTQYEVSDGEGVYNLVNSN 234

Qy 209 -PEDWIEGAAVAADYDYPGVVNDTTKQWFKVAASQEVADKWAERAEVTPVTGKRLMEARV 267
Db 235 DPQKWIHETARLAKVEIGNITNDEIKSHYKNGNNA-----LMQOEVMFPAQVRSLENAQRN 289

Qy 268 TAGYIHLWFDITYN 280
Db 290 TAGFIHLWFKTFV 302

OTHER INFORMATION: Obtained from an environmental sample
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-10-421-654-6

Query Match      79.3%; Score 1206.5; DB 15; Length 280;
Best Local Similarity 76.2%; Pred. No. 8.9e-109;
Matches 215; Conservative 34; Mismatches 28; Indels 5; Gaps 2;

Qy 1 MKKKVLALAAAMVALAAPVQS--VVFQATNNSESPAPILRWSAEDKNEGIN 59
Db 1 MKKKLALATASVIALTAIPQSAFAH-----ESDPIALRWSAESVNEGSSHLWVNR 56

Qy 60 DIMSRNTTIVNPNETALLNWRADLENGIYSADYENPYDNTSYASHFYDPDGTGTYIPF 119
Db 57 DIMSQNTTIVKQNETALLNWRNTLEEGIYSADYKPNPYDNTSFASHFYDPDSEKTYIPF 116

Qy 120 AKHAKETGAKYNLAGAYONQDMQAQFFYGLSLHLGLDVNQPMAAANFTNLSYPMGFH 179
Db 117 AKQAKQTGAKYFKLAGEAYQNDLKNAPFYGLSLHLGLDVNQPMAAANFTNISHPPGFH 176

Qy 180 SKYENFVDITKKNYIVSDNSGYNWKNKANPDWIEGAAVAADYDYPGVVNDTTKQWFKVA 239
Db 177 SKYENFVDITKKNYIVSDNSGYNWKNKANPDWIEGAAVAADYDYPGVVNDTTKQWFKVA 236

Qy 240 AVSQEYADKWAERAEVTPVTGKRLMEARVTAQYIHLWFDITYN 281
Db 237 AVSQDSADKWAERAEVTPVTGKRLMEARVTAQYIHLWFDITYN 278

RESULT 5
US-10-421-654-82
; Sequence 82, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; NAME/KEY: SIGNAL
; LOCATION: (1)...(25)
US-10-421-654-82

Query Match      34.3%; Score 522; DB 15; Length 473;
Best Local Similarity 38.7%; Pred. No. 6.7e-42;
Matches 121; Conservative 43; Mismatches 105; Indels 44; Gaps 10;

Qy 1 MKKKVLALAAAMVALAAPVQS--VVFQATNNSESPAPILRWSAEDKNEGIN 49
Db 1 MKKKLCTALVTAISSGVVITPTTEAQACGIGEVKQENQEHK--RVKWSAEHPHHPNES 58

Qy 50 SHLWVNRADIMSRN--TTIVNPNETALLN--EWRADLENGIYSADYENPYDNTST--- 102
Db 59 THLWTRARNAIQIMSRNQKTVQENELQFLNTPYKELFERGLYDADYLDDEFNDGDTG 118

Qy 103 -----YASHFYDPDGTGTYIPPAKHAKE-----TGAKYFNLAGQAYQNDMQQAFFY 149
Db 119 IDGLIKGGWKSHFYDPDTRKNY----KGEETPTALSQGDQKFKLAGDYFKKEDQQAFFY 174

Qy 150 LGLSLHLYLGDVNPQMAAANFTNL--SYPMGFHSKYENFVDITKKNYIVSDNSGYNWKNKAN 208
Db 175 LGVATHYFTDATQPMHAANFTAVDTSAKFKHSAFENYVTTIQTQYEVSDGEGVYNLVNSN 234

Qy 209 -PEDWIEGAAVAADYDYPGVVNDTTKQWFKVAASQEVADKWAERAEVTPVTGKRLMEARV 267
Db 235 DPQKWIHETARLAKVEIGNITNDEIKSHYKNGNNA-----LMQOEVMFPAQVRSLENAQRN 289

Qy 268 TAGYIHLWFDITYN 280
Db 290 TAGFIHLWFKTFV 302
```

```
RESULT 7
US-10-421-654-98
; Sequence 98, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; PRIOR FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Unknown
; OTHER INFORMATION: Obtained from an environmental sample
; NAME/KEY: SIGNAL
; LOCATION: (1)....(25)
US-10-421-654-98

Query Match 33.4%; Score 509; DB 15; Length 473;
Best Local Similarity 37.8%; Pred. No. 1.2e-40;
Matches 119; Conservative 43; Mismatches 109; Indels 44; Gaps 10;

Qy 1 MKKKVLALAAAMVALAAPVQSV-----VFAQTNNSESPAPILRWSAEDKHNEGIN 49
Db 1 MKKKLCTWALVTAISSGVVAIPTVASACGMEVMKQOEHK--RVKWSAEHPHSHES 58
Qy 50 SHLWVNRADIMSRN--TTIVNPNETALLN--EWRADLENGIYSADYENPYDNST--- 102
Db 59 TMLWIAQNAIQIMSRNQDKTQVENLQFLNPEYKELFERGLYDADYLDDEFNDGGTGTIG 118
Qy 103 -----YASHFYDDPTGTYIIPAKHAK-----TGAKYFNLAGQAYQNDMQQAFY 149
Db 119 IDGLIRGCKWKSHPYDPDKKNY----KGESEPTALSQGDYFKLAGEYFKKNDWKQAFY 174
Qy 150 LGLSLHYLDVNPQMHAAFNLTLS--YPMGFPHSKYENFVDTIKNNYIVSDSNGYNNWKGAN 208
Db 175 LGVATHYFTDATQPMHAAFNFTAVDRSAIKFHSFAFEDYVTTIQEQPKVSDGEGKYNLVNSN 234
Qy 209 -PEDWIEGAAVAAKQDYPGVVNDTTKDFVKAASQEVADKWRARVTPVTGKRLMEARV 267
Db 235 DPQKWIHETARLAKVEIGNINDVIKSHYNNKGNNA-----LWQGEVMPAVQVSLEQAQRN 289
Qy 268 TAGYIHLWFDYVNR 282
Db 290 TAGFIHLWFKTYVGK 304

RESULT 8
US-10-421-654-32
; Sequence 32, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; PRIOR FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Unknown
; OTHER INFORMATION: Obtained from an environmental sample
; NAME/KEY: SIGNAL
; LOCATION: (1)....(30)
US-10-421-654-106

Query Match 32.9%; Score 501; DB 15; Length 473;
Best Local Similarity 36.5%; Pred. No. 7.4e-40;
Matches 115; Conservative 49; Mismatches 107; Indels 44; Gaps 10;

Qy 1 MKKKVLALAAAMVALAAPVQSV-----VFAQTNNSESPAPILRWSAEDKHNEGIN 49
Db 1 MKKKLCTWALVTAISSGVVAIPTVASACGMEVMKQOEHK--RVKWSAEHPHSHES 58
Qy 50 SHLWVNRADIMSRN--TTIVNPNETALLN--EWRADLENGIYSADYENPYDNST--- 102
Db 59 TMLWIAQNAIQIMSRNQDKTQVENLQFLNPEYKELFERGLYDADYLDDEFNDGGTGTIG 118
Qy 103 -----YASHFYDDPTGTYIIPAKHAK-----TGAKYFNLAGQAYQNDMQQAFY 149
Db 119 IDGLIRGCKWKSHPYDPDKKNY----KGESEPTALSQGDYFKLAGEYFKKNDWKQAFY 174
Qy 150 LGLSLHYLDVNPQMHAAFNLTLS--YPMGFPHSKYENFVDTIKNNYIVSDSNGYNNWKGAN 207
Db 175 LGVATHYFTDATQPMHAAFNFTAVDRSAIKFHSFAFENYVTTVQTFEVKDDKGTYNLVNSD 234
Qy 208 NPEDWIEGAAVAAKQDYPGVVNDTTKDFVKAASQEVADKWRARVTPVTGKRLMEARV 267
Db 235 DPQKWIHETARLAKVEIGNINDVIKSHYNNKGNNA-----LWQGEVMPAVQVSLEQAQRN 289
Qy 268 TAGYIHLWFDYVNR 282
Db 290 TAGFIHLWFKTYVGK 304

RESULT 9
US-10-421-654-106
; Sequence 106, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Unknown
; OTHER INFORMATION: Obtained from an environmental sample
; NAME/KEY: SIGNAL
; LOCATION: (1)....(30)
US-10-421-654-106
```

```

RESULT 11
US-10-424-599-239481
; Sequence 239481, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 239481
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURES:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_58277C.1.pep
US-10-424-599-239481

Query Match          6.7%; Score 102.5; DB 12; Length 272;
Best Local Similarity 21.8%; Pred. No. 0.19;
Matches 51; Conservative 37; Mismatches 85; Indels 61; Gaps 13;

Qy      25  QTNNSPAPILRWASAEKNEGINSHLWIVNRAIDIMSRNTITVNPETALLNEWADL 84
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      53  EVNKPEQPPPI---PTENSYGLYGHHDENVQV-----PSITTNPTS---YHPKTEF 100

Qy      85  ENGIYSADYENPYDNSTYASHP-----YBPDGTQTY-IPPAKIAKE-----TGAKYF 131
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     101  ENS-----NKKYNNNDAYNTRFSETGYNNNKDSYEGNQVLSDTKYTEEEYNNNNKYH 154

Qy     132  NLAGQAYQNDMDQAPPYGLSLSLHYLDGVNQPMHAANFTNLSYPMGFHSKYVENFVDTIKN 191
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     155  N-----SQVNNQK-----YYNDAANGIYSYNNNNNY-----NANNRYNTYNN 195

Qy     192  NYIVSDSGSYNNMKGANPEDWIEGAA-----VAAKQDYPGVVNDT-----TDWPF 236
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     196  NNAVNGYNG--ERCGMDSDFLEGGKYFHDVAAEKYNPFTNYGDSRSRVNTNNWY 247

```


Search completed: July 27, 2004, 14:49:15
Job time : 48 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2004, 14:38:52 ; Search time 17 Seconds
(without alignments)
1595.648 Million cell updates/sec

Title: PCT-US03-12556-2
Perfect score: 1522
Sequence: 1 MKKKVLALAMVALAAPVQS.....EAQRTAGYIHLNFDFTYVNR 282

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252.5	82.3	592	2 PS0197	phospholipase C (E
2	511.5	33.6	289	2 C43868	lecithinase - List
3	511	33.6	289	2 AF1100	phospholipase C (i
4	509	33.4	289	2 S20888	phospholipase C (E
5	314.5	20.7	178	2 T43165	probable phospholi
6	254	16.7	398	2 B30565	phospholipase C (E
7	231	15.2	398	2 JQ0366	phospholipase C (E
8	229	15.0	398	2 A30565	phospholipase C (E
9	227	14.9	398	2 B49231	phospholipase C, a
10	109.5	7.2	245	2 H97025	phospholipase C re
11	108.5	7.1	2140	2 F95074	serine proteinase,
12	107	7.0	2144	2 A97942	metalloproteinase,
13	103	6.8	625	2 G97076	alkaline phosphata
14	102.5	6.7	470	2 AF1972	hypothetical prote
15	102	6.7	562	2 S55516	glucan 1,3-beta-gl
16	102	6.7	2358	2 T39569	probable alpha-glu
17	99	6.5	273	2 D82862	conjugal transfer
18	98	6.4	472	2 AE0620	hypothetical prote
19	97.5	6.4	363	2 E9016	outer membrane pro
20	97	6.4	468	2 H64802	ybfM protein - Esc
21	97	6.4	653	2 H85620	partial fimbrial u
22	97	6.4	653	2 F90756	partial fimbrial u
23	96.5	6.3	532	2 D96966	beta-mannanase Man
24	95.5	6.3	418	1 A53888	thermolabile hemol
25	95.5	6.3	1283	2 S52500	oxyesterol-binding
26	95	6.2	441	2 E90051	hypothetical prote
27	95	6.2	624	2 E64483	hypothetical prote
28	95	6.2	1599	2 S22737	glucosyltransferase
29	94.5	6.2	305	2 H86245	bifunctional nucle

30	94	6.2	623	2 B83399	quinoprotein alcoh
31	94	6.2	837	1 A31842	endo-1,4-beta-xyla
32	93.5	6.1	277	1 S04072	fimbrial adhesin F
33	93.5	6.1	770	2 F82383	methyl-accepting c
34	93.5	6.1	787	2 G81692	inner membrane pro
35	93	6.1	405	2 S73853	hypothetical prote
36	92.5	6.1	304	2 S24263	seed storage prote
37	92.5	6.1	538	2 H87347	hypothetical prote
38	92.5	6.1	910	1 S73361	dnaj homolog prote
39	92	6.0	782	2 A82940	hypothetical prote
40	91.5	6.0	358	2 AE2218	hypothetical prote
41	91.5	6.0	363	2 S43159	outer membrane por
42	91.5	6.0	441	2 G83212	hypothetical prote
43	91.5	6.0	474	1 JC6042	cytosol nonspecific
44	91	6.0	296	2 T35345	chitinase - Strept
45	91	6.0	342	2 A87389	hypothetical prote

ALIGNMENTS

RESULT 1

PS0197
phospholipase C (EC 3.1.4.3) precursor - Bacillus cereus
N:Alternate names: cereolysin AB; phosphatidylcholine-hydrolyzing phospholipase C
C:Species: Bacillus cereus
C:Date: 30-Sep-1991 #sequence_revision 26-Apr-1996 #text_change 15-Oct-1999
C:Accession: S18978; S01949; A32042; S03200; I40161
R:Kuzmin, N.P.; Gavrilenko, I.V.; Krukov, V.M.; Karpov, A.V.
submitted to the EMBL Data Library January 1992
A:Description: Fused genes phospholipase C and sphingomyelinase Bacillus cereus ensuring
A:Reference number: S18978
A:Accession: S18978
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-592 <KUZ>
A:Cross-references: EMBL:X64140; NID:g39421; PIDN:CAA45501.1; PID:g39422
R:Ikedo, K.; Inoue, S.; Anasaki, C.; Teshima, K.; Ikezawa, H.
J. Biochem. 110, 88-95, 1991
A:Title: Kinetics of the hydrolysis of monodispersed and micellar phosphatidylcholines c
A:Reference number: PS0197; MUID:92041771; PMID:1939031
A:Accession: PS0197
A:Molecule type: protein
A:Residues: 39-49 <IKS>
R:Gavrilenko, I.V.; Baida, G.E.; Karpov, A.V.; Kuz'min, N.P.
Bioorg. Khim. 19, 133-138, 1993
A:Title: [Nucleotide sequence of phospholipase C and sphingomyelinase genes from Bacillu
A:Reference number: I40160; MUID:93249510; PMID:8387306
A:Accession: I40160
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-592 <RES>
A:Cross-references: EMBL:X64140; NID:g39421; PIDN:CAA45501.1; PID:g39422; EMBL:X64141; N
R:Johansen, T.; Haugli, F.B.; Ikezawa, H.; Little, C.
Nucleic Acids Res. 16, 10370, 1988
A:Title: Bacillus cereus strain SE-1: nucleotide sequence of the sphingomyelinase C gene
A:Reference number: S01949; MUID:89057484; PMID:2848222
A:Accession: S01949
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-283 <JOH>
A:Cross-references: EMBL:X12854; NID:g39444; PIDN:CAA31332.1; PID:g295765
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1988
R:Gilmore, M.S.; Cruz-Rodaz, A.L.; Leimeister-Waechter, M.; Kreft, J.; Goebel, W.
J. Bacteriol. 171, 744-753, 1989
A:Title: A Bacillus cereus cytolytic determinant, cereolysin AB, which comprises the pho
A:Reference number: A32042; MUID:89123149; PMID:2536680
A:Accession: A32042
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-8, 'G', 10-13, 'L', 15-30, 'QRFQVIP', 38-76, 'L', 78-119, 'Y', 121-211, 'D', 213-225, '
A:Cross-references: GB:M4149; NID:g142672; PIDN:AA91819.1; PID:g142673
R:Yamada, A.; Tsukagoshi, N.; Ueda, S.; Sasaki, T.; Makino, S.; Nakamura, S.; Little, C.

Eur. J. Biochem. 175, 213-220, 1988
A:Title: Nucleotide sequence and expression in *Escherichia coli* of the gene coding for s
A:Reference number: S01130; MUID:88296483; PMID:2841128
A:Accession: S03200
A:Molecule type: DNA
A:Residues: 166-211, 'D', 213-225, 'A', 227-238, 'R', 240-281, 'N', 283 <YAM>
A:Cross-references: EMBL:X12711; NID:g39439; PIDN:CAA31213.1; PID:g57735
C:Genetics:
A:Gene: cerAB; plc
C:Superfamily: Bacillus cereus phospholipase C
C:Keywords: phosphoric diester hydrolase
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-38/Domain: propeptide #status predicted <PRO>
F:39-283/Product: phospholipase C #status predicted <MAT>

Query Match 82.3%; Score 1252.5; DB 2; Length 592;
Best Local Similarity 80.2%; Pred. No. 3.2e-91;
Matches 227; Conservative 23; Mismatches 32; Indels 1; Gaps 1;

Qy 1 MKKKVLAAAMVLAAPVQSVVFAQTNNSESPAPIL-RWSAEDKHNGINSHLMTVNRAI 59
Db 1 MKKKVLAAALAITVVAPLQSVAFAPAHENOGGKIKIVHWSAEDKHGKGVNHLMTVNRAI 60

Qy 60 DIMSNTTIVNPNETALLNEWRADLENGIYSADYENPYDSTYASHFYDPTGTGYTYPF 119
Db 61 DIMSENTTLVKQDRVAQLNWRTELENGIYAADYENPYDSTFASHFYDPDNGKTYTYPF 120

Qy 120 AKHAKETGAKYFNLAGQYQNDMQQAFYGLSLHYLGDVNPMAAANFNLSYPMGFH 179
Db 121 AKQAKETGAKYFKLAGESYKNDKMQAFYGLSLHYLGDVNPMAAANFNLSYPMGFH 180

Qy 180 SKYENFVDTIKNNYIVSDSGYNNWKGANPEDMIEGAAVAAKQDYPGVVNDTTKDMFYKA 239
Db 181 SKYENFVDTIKNNYKVTGNGYNNWKGANPERIHGAAVAAKQDYSGLVNDTTKDMFYKA 240

Qy 240 AVSQYADKWAERAEVTPVTKRLMEAQRTAGYIHLWFDTYNVR 282
Db 241 AVSQYADKWAERAEVTPMTGKRLMDAQRTAGYIQLWFDTYGDR 283

RESULT 2
C43868
lecithinase - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1998
C:Accession: C43868
R:Vazquez-Boland, J.A.; Kocks, C.; Dramsi, S.; Ohayon, H.; Geoffroy, C.; Mengaud, J.; Co
Infect. Immun. 60, 219-230, 1992
A:Title: Nucleotide sequence of the lecithinase operon of *Listeria monocytogenes* and pos
A:Reference number: A43868; MUID:92104678; PMID:1309513
A:Accession: C43868
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <VAZ>
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:74437, NCBIIP:74462)
C:Superfamily: Bacillus cereus phospholipase C

Query Match 33.6%; Score 511.5; DB 2; Length 289;
Best Local Similarity 38.7%; Pred. No. 4.5e-33;
Matches 98; Conservative 52; Mismatches 96; Indels 7; Gaps 2;

Qy 24 AQTNNSESPAP-ILWSAEDKHNGINSHLMTVNRADIMSRNTTIVNPNETALLNEWRA 82
Db 37 ARPHDIDSKLPHKLSWSADNPTNTDVNTHYLWLFKQAEKILAKDVNHRANLNLKPKDK 96

Qy 83 DLENGIYSADYENPYDSTYASHFYDPTGTGYTYPFPAKHAKETGAKYFNLAGQYQND 142
Db 97 QIAQIYDADHKNPYDSTYSLSHFNPDRDNTYLPGFANAKITGAKYFNQSVTDYREGK 156

Qy 143 MQQAFYGLSLHYLGDVNPMAAANFNLSYPMGFHSHKYENFVDTIKNNYIVSDSGYNN 202
Db 157 FDTAFYKGLGLAIHYTDTISQPMHANNFTAISYPPGCHAYENYVDTIKHNYQATDWMVAK 216

Qy 203 NMGKANPEDMIEGAAVAAKQDYPGVVNDTTKDMFYKAASQYADKWAERAEVTPVTKRLM 262
Db 217 RFGSDDDKDWLYENAKRAKADYPKIVNAKTKSYLVG-----NSEWKKDVTVEFTGARLR 270

Qy 263 EAQRTVAGYIHLW 275
Db 271 DSQQTLAGFLEFW 283

RESULT 3

AF1100
phospholipase C [imported] - *Listeria monocytogenes* (strain BGD-e)
C:Species: *Listeria monocytogenes*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AF1100
R:Glaser, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of *Listeria species*.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1100
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00732.1; PID:g16409570; GSPDB:GN00177
A:Experimental source: strain BGD-e
C:Genetics:
A:Gene: plcB
C:Superfamily: Bacillus cereus phospholipase C

Query Match 33.6%; Score 511; DB 2; Length 289;
Best Local Similarity 37.1%; Pred. No. 5e-33;
Matches 106; Conservative 55; Mismatches 107; Indels 18; Gaps 5;

Qy 2 KKKVLALAAVVA-LAAPVQSVV-----FAQT---NNSESPAP-ILWSAEDKHNGIN 49
Db 4 KKVVLGMLIASLVLPVPTIKANACCDEYLOTPAAPHDIDSKLPHKLSWSADNPTNTDVN 63

Qy 50 SHLWIVNRADIMSRNTTIVNPNETALLNEWRADLENGIYSADYENPYDSTYASHFYD 109
Db 64 THYLFKQAEKILAKDVNHRANLNLKPKDKQIAQGIYDADHKNPYDSTYSLSHFYN 123

Qy 110 PDTGTTTYPFAKHAKETGAKYFNLAGQYQNDMQQAFYGLSLHYLGDVNPMAAANF 169
Db 124 PDRDNTYLPGFANAKITGAKYFNQSVTDYREGKFTAPYKGLGLAIHYTDTISQPMHANNF 183

Qy 170 TNLSPMGFHSKYENFVDTIKNNYIVSDSGYNNWKGANPEDMIEGAAVAAKQDYPGVVN 229
Db 184 TAISYPPGCHAYENYVDTIKHNYQATDWMVAKPCSDDDKDWLYENAKRAKADYPKIVN 243

Qy 230 DTTKDMFYKAASQYADKWAERAEVTPVTKRLMEAQRTAGYIHLW 275
Db 244 AKTKSYLVG-----NSEWKKDVTVEFTGARLRDSQQTLAGFLEFW 283

RESULT 4

S20888
phospholipase C (EC 3.1.4.3) precursor - *Listeria monocytogenes* (fragment)
C:Species: *Listeria monocytogenes*
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 24-Sep-1999
C:Accession: S20888
R:Donann, E.; Wehland, J.; Rohde, M.; Pistor, S.; Hartl, M.; Goebel, W.; Leimeister-Wae
EMBO J. 11, 1981-1990, 1992
A:Title: A novel bacterial virulence gene in *Listeria monocytogenes* required for host ce
A:Reference number: S20887; MUID:92258410; PMID:1582425
A:Accession: S20888
A:Molecule type: DNA
A:Residues: 1-289 <DOM>
A:Cross-references: EMBL:X59723; NID:g44091; PIDN:CAA42408.1; PID:g444093

Db 103 NNNFSKNSWYLAYSIPDTGESQIRKFSALARYEWRGNKYKQATFYLGEMAHYFGDIDTP 162
QY 164 MHAANFTNLSYPMGPHSKYENFVDTIKNNYIVSDSNGYWNKGANPEMDWIEGAQVAAKQD 223
Db 163 YHPANVT--AVDSAGHVKFETFAERKEQYKINTVGCCTN-----EDFY--ADILKNKD 212
QY 224 YPGVVNDTTKDW---FVKAASQBYA-----DKW--RAEVTPTVKRLMEARQVTAQ 270
Db 213 F----NAWSKEYARGFAGTKGSIYSHASHMSHSDWDMDYAAKVT-----LANSQKGTAG 262
QY 271 YIH 273
Db 263 YIY 265

RESULT 8

A30565
phospholipase C (EC 3.1.4.3) precursor - Clostridium perfringens
N/Alternate names: alpha toxin
C/Species: Clostridium perfringens
C/Date: 23-Mar-1989 #sequence_revision 23-Mar-1989 #text_change 21-Jul-2000
C/Accession: A30565; A32418; A30564; S06996
R/TSO, J.Y.; Siebel, C.
Infect. Immun. 57, 468-476, 1989
A/Title: Cloning and expression of the phospholipase C gene from Clostridium perfringens
A/Reference number: A30565; MUID:89108588; PMID:2536356
A/Accession: A30565
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-398 <TSO>
R/Okabe, A.; Shimizu, T.; Hayashi, H.
Biochem. Biophys. Res. Commun. 160, 33-39, 1989
A/Title: Cloning and sequencing of a phospholipase C gene of Clostridium perfringens.
A/Reference number: A32418; MUID:89228034; PMID:2540749
A/Accession: A32418
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-16,'S',18-398 <OKA>
A/Cross-references: GB:M24904; NID:G144889; PIDN:AAA23272.1; PID:G144889
R/Ticball, R.W.; Hunter, S.E.C.; Martin, K.L.; Morris, B.C.; Shuttleworth, A.D.; Rubidge
Mol. Microbiol. 3, 383-392, 1989
A/Title: Molecular cloning and nucleotide sequence of the alpha-toxin (phospholipase C)
A/Reference number: A30564; MUID:89108574; PMID:2536355
A/Accession: A30564
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-16,'S',18-398 <TIT>
R/Leslie, D.; Fairweather, N.; Pickard, D.; Dougan, G.; Kehoe, M.
Mol. Microbiol. 3, 383-392, 1989
A/Title: Phospholipase C and haemolytic activities of Clostridium perfringens alpha-toxin
A/Reference number: S06996; MUID:89313290; PMID:2546005
A/Accession: S06996
A/Molecule type: DNA
A/Residues: 1-16,'S',18-398 <LES>
A/Cross-references: EMBL:X13608; NID:G40554; PIDN:CAA31943.1; PID:G40555
C/Genetics:
A/Gene: cpa
C/Keywords: phosphoric diester hydrolase
F/1-28/Domain: signal sequence #status predicted <SIG>
F/29-398/Product: phospholipase C #status predicted <MAT>

Query Match 15.0%; Score 229; DB 2; Length 398;
Best Local Similarity 25.8%; Pred. No. 1.5e-10;
Matches 79; Conservative 44; Mismatches 103; Indels 74; Gaps 13;

QY 1 MKKKYLALAAWVALA----APQSVVFAQTNNSSPAPILRWSAEDKHNEGINSHLMTVN 56
Db 1 MKRKICKALICAAATLTMAGASTKYA-----W---DGKIDGTGTHAMIVT 44
QY 57 RAIDMSRNTTIVNPNET-----ALLNWRADLENGIYSADYENPYDNTSYASHFYDPT 112
Db 45 QGVSIENDLSKNEPESVRKNLEILKENMHQLQGSTPYDKNAYD--LYQDHFWDPT 102

QY 113 GTTYIP-----FAKHAKETG-----AKYFNLAGQAYQNDMQQAFYFGLSLHLGLDGNQP 163
Db 103 DNNFSKNSWYLAYSIPDTGESQIRKFSALARYEWRGNKYKQATFYLGEMAHYFGDIDTP 162
QY 164 MHAANFTNLSYPMGPHSKYENFVDTIKNNYIVSDSNGYWN-----WGAN 208
Db 163 YHPANVT--AVDSAGHVKFETFAERKEQYKINTAGCKTNEAFYTDILKNKDFNAWS--- 217
QY 209 PEDWIEGAANAQKQDYPCVVNDTTKDMFVKAASQBYA---KWRAEVTPVTVKRLMEARQV 267
Db 218 -KEYARGFAGTKGSIY-----YSHASHMSHSDWDMDYAAKVT-----LANSQKGTAG 259
QY 268 TAGYIH 273
Db 260 TAGIY 265

RESULT 9

B49231
phospholipase C, alpha-toxin - Clostridium perfringens
C/Species: Clostridium perfringens
C/Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C/Accession: B49231
R/Katayama, S.; Matsushita, O.; Minami, J.; Mizobuchi, S.; Okabe, A.
Infect. Immun. 61, 457-463, 1993
A/Title: Comparison of the alpha-toxin genes of Clostridium perfringens type A and C strains
A/Reference number: B49231; MUID:93138764; PMID:8423073
A/Contents: type C, NCIB 10662
A/Accession: B49231
A/Status: preliminary
A/Molecule type: nucleic acid
A/Residues: 1-398 <KAT>
A/Note: sequence extracted from NCBI backbone (NCBI:123169, NCBI:123171)

Query Match 14.9%; Score 227; DB 2; Length 398;
Best Local Similarity 26.7%; Pred. No. 2.1e-10;
Matches 81; Conservative 46; Mismatches 108; Indels 68; Gaps 15;

QY 1 MKKKV----LALAAWVALAAPQSVVFAQTNNSSPAPILRWSAEDKHNEGINSHLMTVN 56
Db 1 MKRKICKALICRTATSLMAGASTKYA-----W---DGKIDGTGTHAMIVT 44
QY 57 RAIDMSRNTTIVNPNET-----ALLNWRADLENGIYSADYENPYDNTSYASHFYDPT 112
Db 45 QGVSIENDLSKNEPESVRKNLEILKENMHQLQGSTPYDKNAYD--LYQDHFWDPT 102
QY 113 GTTYIP-----FAKHAKETG-----AKYFNLAGQAYQNDMQQAFYFGLSLHLGLDGNQP 163
Db 103 DNNFSKNSWYLAYSIPDTGESQIRKFSALARYEWRGNKYKQATFYLGEMAHYFGDIDTP 162
QY 164 MHAANFTNLSYPMGPHSKYENFVDTIKNNYIVSDSNGYWNKGANPEMDWIEGAQVAAKQD 223
Db 163 YHPANVT--AVDSAGHVKFETFAERKEQYKINTAGCKTN-----EDFY--ADILKNKD 212
QY 224 YPGVVNDTTKDW---FVKAASQBYA-----DKW--RAEVTPTVKRLMEARQVTAQ 270
Db 213 F----NAWSKEYARGFAGTKGSIYSHASHMSHSDWDMDYAAKVT-----LANSQKGTAG 262
QY 271 YIH 273
Db 263 YIY 265

RESULT 10

H97025
phospholipase C related protein [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C/Accession: H97025
R/Nolling, J.; Brston, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum

Qy 143 -----MQAPFFYLGL--SLHYLGDVNPQPMHAAPF-----TNLSYPMGPF 178
 Db 179 EGISSEVERAQVQPMNMHARKEIGVEAIDYLSINAPF--GKNPFDGRGMWISNIDTGTDTY 237
 Qy 179 HSKYENFVDTIK-----NNYIVSD-----SNGYWNWKGANPDWTGGAATAA 220
 Db 238 RHKAMRIDDDAKASMRPFKKEDLKGTDKNYWLSDKIPHAFNYNGGKITVEKYDDG----- 292
 Qy 221 KODY-----PGVV--NDTTKD 234
 Db 293 -RDYFDPHGMHGIAGILAGNDTEQD 315

RESULT 12
 A97942
 metalloproteinase (EC 3.4.21.-) A [imported] - Streptococcus pneumoniae (strain
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 C:Accession: A97942
 R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McArthur,
 Y, P.; Sun, P.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskun
 A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A;Reference number: A97872; MUID:21429245; PMID:11544234
 A;Accession: A97942
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-2144 <R>
 A;Cross-references: GB:AE007317; PIDN:AAK99365.1; PID:g15458138; GSPDB:GN00174
 C;Genetics:
 A;Gene: prtA
 C;Keywords: hydrolase; serine proteinase

Query Match 7.0%; Score 107; DB 2; Length 2144;
 Best Local Similarity 23.1%; Pred. No. 5.7;
 Matches 76; Conservative 36; Mismatches 113; Indels 104; Gaps 19;

Qy 1 MKKKV--LALAAMVALAAPVQSVFPAQTNNSSPAPI--LRWSAEDKHNEGINSHLWI- 54
 Db 1 MKKSTVLSLTAAVILAAAPNEVLADTSSSEDALSIDSKKVVVDKETENKEKKHDIH 60

Qy 55 --VNRAIDMSRNTTV-----NP-----NETALLNEWRADLENGIYSADY---- 93
 Db 61 NAIETSKDTEEKTITIEEKVEVSKNPVIDTKTSNEEAKIKEENSQSGDHTDSFVNKN 120

Qy 94 -ENP-YDINSTYASHFVDPDTGTYTTPFAKHAKETGAKY-----FNLGAQAYQ-----NQD 142
 Db 121 TENPKKEDKVYTAEPDKESGEKAIKGLSNLKNTKVLYTYDRIFN--GSAIETTPDNL 178

Qy 143 -----MQQAFVILGH--SLHYLDVNPQPMHAAPF-----TNLS 173
 Db 179 KIKQIEGSIERRAQVQPMNMHARKEIGVEAIDYLSINAPF--GKNPFDGRGMWISNID 237

Qy 174 YPMGFHSKVENFVDTIK-----NNYIVSD-----SNGYWNWKGANPDWTGGAATAA 215
 Db 238 TGTDYRHKAMRIDDDAKASMRPFKKEDLKGTDKNYWLSDKIPHAFNYNGGKITVEKYDDG 297

Qy 216 AAVAAKQDY-----PGVV--NDTTKD 234
 Db 298 -----RDYFDPHGMHGIAGILAGNDTEQD 320

RESULT 13
 G97076
 alkaline phosphatase superfamily protein [imported] - Clostridium acetobutylicum
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
 C:Accession: G97076
 R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson,
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001

Search completed: July 27, 2004, 14:47:48
Job time : 18 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2004, 23:41:59 ; Search time 3697 Seconds
(without alignments)
9953.540 Million cell updates/sec

Title: PCT-US03-12556-1

Perfect score: 849

Sequence: 1 atgaaaagaagattagc.....atagctatgtaaatcgctaa 849

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.ats.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.ats.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	559.2	65.9	1675	1	BTY16268	Y16268 Bacillus th
2	556	65.5	2245	1	BACCRAB	M24149 Bacillus ce
3	556	65.5	2349	1	BGNE	X6141 B. cereus ge
4	556	65.5	2700	1	BCPLCSMC	X12854 B. cereus g
5	554	65.3	2201	1	BCFUSE	X6140 B. cereus fu
6	552.8	65.1	290214	1	AE017026	AE017026 Bacillus
7	552.8	65.1	301289	1	AE017000	AE017000 Bacillus
8	546.4	64.4	2130	1	AY195600	AY195600 Bacillus
9	255.6	30.1	1622	1	BCPLC	X12711 B. cereus g
10	189.4	22.3	280	1	BACHEMOL	M3411 B. cereus he
11	176.8	20.8	825	1	AY150839	AY150839 Listeria
12	154.4	18.2	870	6	AX413749	AX413749 Sequence
13	154.4	18.2	870	6	AX415655	AX415655 Sequence
14	154.4	18.2	5648	1	LISACTLDH	M82881 Listeria mo
15	154.4	18.2	160050	1	AL591974	AL591974 Listeria
16	154.4	18.2	34980	6	AX638811	AX638811 Sequence
17	152.8	18.0	3031	1	LMACTPLC	X59723 L.monocytog
18	148.8	17.5	758	1	AY150836	AY150836 Listeria
19	148.2	17.5	758	1	AY150830	AY150830 Listeria
20	148.2	17.5	758	1	AY150840	AY150840 Listeria
21	148.2	17.5	758	1	AY150841	AY150841 Listeria
22	145.6	17.1	758	1	AY150837	AY150837 Listeria
23	144.6	17.0	758	1	AY150835	AY150835 Listeria
24	144	17.0	758	1	AY150833	AY150833 Listeria
25	144	17.0	758	1	AY150834	AY150834 Listeria
26	137.4	16.2	758	1	AY150838	AY150838 Listeria
27	118.2	13.9	3417	1	LIV249805	AJ249805 Listeria
28	117.4	13.8	4097	1	LIU19035	U19035 Listeria iv
29	117.2	13.8	740	1	AY150832	AY150832 Listeria
30	64.2	7.6	701	1	AY150831	AY150831 Listeria
31	57.4	6.8	4562	6	AX416781	AX416781 Sequence
32	57.2	6.7	1237	1	AF072123	AF072123 Clostridi
33	56.8	6.7	250029	3	AE014820	AE014820 Plasmodi
34	55	6.5	125623	3	AC115599	AC115599 Dictyoste
35	55	6.5	183357	2	BX569779	BX569779 Danio rer
36	54.2	6.4	340552	3	PFA929354	AL929354 Plasmodi
37	54	6.4	1234	1	AB061869	AB061869 Clostridi
38	53.4	6.3	156323	2	BX296532	BX296532 Danio rer
39	52.6	6.2	33321	3	AC116986	AC116986 Dictyoste
40	51	6.0	110000	3	AC116984_3	Continuation (4 of
41	50.8	6.0	84618	5	AL591365	AL591365 Zebrafish
42	50.6	6.0	110000	2	PFMAL13_24	Continuation (25 o
43	50.2	5.9	192578	5	AL732455	AL732455 Zebrafish
44	49.8	5.9	256879	3	AC116982	AC116982 Dictyoste
45	49.4	5.8	49144	2	AC131380	AC131380 Strongylo

ALIGNMENTS

RESULT 1
BTY16268
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

BTY16268 1675 bp DNA linear BCT 02-OCT-1998
Bacillus thuringiensis plca gene and partial plcB gene
Y16268
Y16268.1 GI:2815226
phospholipase C; plcA gene; plcB gene; sphingomyelinase.
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.

1
Lougren, A., Carlson, C.R., Eskils, K. and Kolsto, A.B.
Localization of putative virulence genes on a physical map of the
Bacillus thuringiensis subsp. gelechiae chromosome

JOURNAL	Curr. Microbiol. 37 (4), 245-250 (1998)
PUBMED	9732531
REFERENCE	2 (bases 1 to 1675)
AUTHORS	Lovgren,A.
TITLE	Direct Submission
JOURNAL	Submitted (21-JAN-1998) A. Lovgren, Dept of Microbiology, Stockholm
FEATURES	University, S-106 91 Stockholm 106 91, SWEDEN
SOURCE	Location/Qualifiers
	1..1675
	/organism="Bacillus thuringiensis"
	/mol_type="genomic DNA"
	/strain="Bt13"
	/sub_species="gelechiaie"
	/db_xref="taxon:1428"
	/clone="PLC9:51"
gene	484..1335
	/gene="plcA"
CDS	484..1335
	/gene="plcA"
	/codon_start=1
	/transl_table=11
	/product="phosphatidyl-degrading phospholipase C"
	/protein_id="CAA76148.1"
	/db_xref="GI:2815227"
	/db_xref="GOA:O52864"
	/db_xref="SPTREMBL:O52864"
	/translation="MKKKVLALAAITLVAPLQSVAFHENDGGORFGVIRWSAEDK
	HKEGVNHLWLVNRAIDMSRNTLVKQDRVALLNWRTELENGIYAADYENPYDNS
	TFASHFYDPNGKTYIPYAKOAKETGAKYFKLAGESYKQKMKQAFYFLGLSLYLGD
	VNQPMHAAFNGLSYPOGFHSKYENFVDITIKDNYKVTGNGYMWKNGTNPEDVHGA
	VVAQDVAGIVNDNTKFWRAVSVQEVADKWRAEVTPMTGKRLMDAQRTVAGYIQLW
	FDTYGNR"
gene	1529..1675
	/gene="plcB"
CDS	1529..1675
	/gene="plcB"
	/codon_start=1
	/transl_table=11
	/product="sphingomyelinase"
	/protein_id="CAA76149.1"
	/db_xref="GI:2815228"
	/db_xref="SPTREMBL:O52865"
	/translation="MTHNVYMLSTNLYPNWQTERADLIGAADYIKNQDVVILNEVFD
	NSASD"
ORIGIN	
Query Match	65.9%; Score 559.2; DB 1; Length 1675;
Best Local Similarity	79.3%; Pred. No. 2.5e-105;
Matches	676; Conservative 0; Mismatches 173; Indels 3; Gaps 1;
Qy	1 ATGAAAAGAAAGTATTAGCACTAGCAGCTATGCTGCTTGTAGCTGCGCCAGTTCAAAGT 60
Db	484 ATGAAAAGAAAGTACTTCTTTAGCGGAGCTATTACATTAGTTGCTCCATTACAAAGT 543
Qy	61 GTAGTATTGTCACAAACAATAATAGTG---AAAGTCTCGCACCACTTTTAAGATGGTCA 117
Db	544 GTTGCTATTGCTCATGAATATGATGGGGAGAGAGATTGGAGTTATTCGGCGTGGTCT 603
Qy	118 GCTGAGGATAGCATATAATAGGGGATTAATCTCTCATTTGTGGATTGTAATCGTGCATAT 177
Db	604 GCTGAAGATAAACAATAAAGAGCGGTGAATCTCATTTATGATGTTGTAACCGTGCAT 663
Qy	178 GACATCATGCTCGTAATACACAGATTGCTGATCCGATGAACCTGATTTAATATGAG 237
Db	664 GATATTATGCTCGTAATACACACTTGTAAACAAGATCGAGTTGCACTATTAAATGAA 723
Qy	238 TGGCGTGCATTTAGAAAATGGTATTATTCTGCTGATTACGAGATCCCTTATTATGAT 297
Db	724 TGGCGTACTGAGTATAGAGACGGTATTATTGCTGCTGACTATGAAATCCCTTATTATGAT 783
Qy	288 AATAGTACATATGCTTCTCATTATGATCCGGATACCTGGAACAACATATATTCCTTTT 357
Db	784 AATAGCACATTTGCTTCACTTTCTATGACCTCGCAATATGGGAAAACCTTATATTCCTGAT 843

31

756

Qy	358 GCGAAACATGCAAAAGAAACAGGCGCAAAATATTTTAACTTCTGCTGCTCAAGCATACCAA 417
Db	844 GCAAGCAGGCAAGGAACTGGAGCTAAATATTTTAAATAGCTGGTGGTCTTACAAA 903
Qy	418 AATCAAGATATGCAAGCAATCTTCTTACATTAGATATTCGCTTCAATTTATTTAGGAGAT 477
Db	904 AATAAGATATGAAACAAGCAATCTTCTTATTTAGGATTTCTCTTCAATTTATTTAGGGAT 963
Qy	478 GTGAATCAGCAATGCAATGCAAGCAATCTTACGAATCTTCTTCAATCCAAATGGTTTCCAT 537
Db	964 GTAAACCAACCGATGCAATGCGGCAAACTTTACGAACCTTTTGGTATCCAAAGGATTTCCAT 1023
Qy	538 TCTAAATACGAAAAATTTTGTGATACAAATAAAAAATAACTATATATTTTTCAGATAGCAAT 597
Db	1024 TCTAAATATGAAACTTTTGTAGATACGATAAAAGATATATATAAGTAAACGATGGAAT 1083
Qy	598 GGATATTGGAATTTGGAAGAGCAAAACCCAGAAATTTGGATTTGAAGAGCAGCGGTAGCA 657
Db	1084 GGATATTGGAATTTGGAAGAGTACAAATCCAGAAATTTGGATTTTCATGGAGCGCAGTAGTT 1143
Qy	658 GCTAAACAAGATTTCTGCGTTGTGAACGATACGACAAAGATTTGGTTTGTAAAAGCA 717
Db	1144 GCGAAACAAGATTTACGCTGGCAATTTGTAATGATAATACGAAAGATTTGGTTTGTAAAGCA 1203
Qy	718 GCCGTATCTCAAGAATATCGAGATAAATGCGCTGCGGAAAGTAAACACCGGTGACAGGAAAG 777
Db	1204 GCTGTATCACAAGATATCGAGATAAATGCGCGCTGAAGTTACACCAATGACAGGTAAG 1263
Qy	778 CGTTTAATGGAAGCGCAGCGGTTACAGCTGGTTATATTCATTTTGTGGTTGTATACGTAT 837
Db	1264 CGTTTAATGGAATGACAAACGTTTACTGCTGGATACATTCAGCTTTGGTTTGTATACGTAT 1323
Qy	838 GTAAATCGCTAA 849
Db	1324 GGAATCGTTAA 1335
RESULT 2	
BACERAB	BACERAB
LOCUS	Bacillus cereus cerA and cerB genes, complete cds.
DEFINITION	M24149.1 GI:142672
ACCESSION	2245 bp DNA linear BCT 11-MAR-1996
VERSION	cerA gene; cerB gene.
KEYWORDS	Bacillus cereus
SOURCE	Bacillus cereus
ORGANISM	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE	1 (bases 1 to 2245)
AUTHORS	Gilmore,M.S., Cruz-Rodriguez,A.L., Leimeister-Wachter,M., Kreft,J. and Goebel,W.
TITLE	A Bacillus cereus cytolytic determinant, cereolysin AB, which comprises the phospholipase C and sphingomyelinase genes: nucleotide sequence and genetic linkage
JOURNAL	J. Bacteriol. 171 (2), 744-753 (1989)
MEDLINE	89123149
PUBMED	2536680
COMMENT	Original source text: Bacillus cereus DNA.
FEATURES	Location/Qualifiers
source	1..2245
	/organism="Bacillus cereus"
	/mol_type="genomic DNA"
	/db_xref="taxon:1396"
gene	201..1052
	/gene="cerA"
CDS	201..1052
	/gene="cerA"
	/codon_start=1
	/transl_table=11
	/protein_id="AAA91819.1"
	/db_xref="GI:142673"
	/translation="MKKKVLALGAAITLVAPLQSVAFHENDGGORFGVIRWSAEDK"

HKEGVNSHLWVNRADIMSRNTTLVKQDRVALLNEMWTELENGIYAADYENPYDYS
TFASHFDPNKGTIPIYAKQAKETGAKYFKLAGESYKNDKMQAFYFLGSLHYLGD
VNQPMHAANFNLSYPOGFHSHKYENFVDTIKDNKYVDGNGYWNWKGTPEDWIHGAA
VVAQDYAGI VINDNTKDFVRAAVSQEVADKWRAEVTPMTCKRLMDAQRVTAGYIQLW
FDTVGNR"

gene

1129..2130

CDS

/gene="cerB"

1129..2130

/gene="cerB"

/codon_start=1

/transl_table=11

/protein_id="AAA91820.1"

/db_xref="GI:142674"

/translation="MKGKLLKGVLSGLVGLALYSGTSAQAEASTNQNDTLKVMTHNV
YMLSTNLPVNGQTERLADFGAARDYIKNDVILNEVDFNSASDRLLGLKKEYPNOT
AVLGRSSSEVDKLTIGNYSSTPEDGVAIKWPVIAEKIYQVFAKGGCPGNLKNKGF
VYTKIKNDRFVHVTHLOAEDSMCGKTSVSRVTKOLKEIQDFIKNNI PNNEYVL
IGDMYVNIKAENNDSEYASMEKTLNASVPSYTGHTATWDATNLSAKYNPDSLA
EYLDYIASKDHANPSYIENKVLQPKSPQWTVTSWPKNIRIWLIIIIQVEATISMK"

ORIGIN

Query Match 65.5%; Score 556; DB 1; Length 2245;

Best Local Similarity 79.1%; Pred. No. 1.1e-104;

Matches 674; Conservative 0; Mismatches 175; Indels 3; Gaps 1;

Qy	1	ATGAAAAAGAAAGTATTAGCACTAGCAGCTATGGTGTCTTGTAGCTGCGCCAGTTCAAAGT	60
Db	201	ATGAAAAAGAAAGTACTTGTCTTAGCGCAGCTATTACATTAGTTGCTCCATTACAAAGT	260
Qy	61	GTAGTATTGTCACAAACAATAATAGTG---AAGTCTCGACCCGATTTTAAGATGGTCA	117
Db	261	GTTCGATTGTCTCATGAAATGATGGGGACAGAGATTGGAGTTATTCCGCGCTGGTCT	320
Qy	118	GCTGAGGATAAGCATAAATGAGGGGATTAACTCTCATTTGTGGATTGTAATCGTCAATT	177
Db	321	GCTGAGATTAACATTAAGAGCGCTGATCTCTCATTTATGATTTGTAACCGTGCAATT	380
Qy	178	GACATCATGTCTCGTAATACACAGTGTGTAATCCGAATGAACATGCATTATTAATGAG	237
Db	381	GATATTATGTCTCGTAATACACACTTGTMAAACAAAGATCGAGTTGCACATTATAATGAA	440
Qy	238	TGGCGTGTGATTAGAAAATGGTATTATTCTCTGCTGATTACGAGAACTCTTATTATGAT	297
Db	441	TGGCGTACTGAGTTAGAGAACGGTATTATGCTGCTGACTATGAAATCTTATTATGAT	500
Qy	298	AATAGTACATGCTCTCTCACTTTATGATCCGGATCTCGGAACAAATATATCTCTTTT	357
Db	501	AATAGCACATTTGCTTTCACATTTCTATGACCTGACATGGGAAACTTATTATCCGTAT	560
Qy	358	GCGAAACATGCAAAAGAACAGCGGCAAAATATTTTAACTTGTGTGTCAGCATACCAA	417
Db	561	GCAAAGCAGGCAAAAGGAAACTGGAGCTAAATATTTTAAATAGCTGGTGAGTCTTTACAA	620
Qy	418	AATCAAGATATGCAGCAAGCATCTCTTACTTAGGATTTATCGCTCATATTATTAGAGAT	477
Db	621	AATAAGATATGAAACAAAGCATCTCTTATTAGGATTTATCTCTTCAATTATTAGGGAT	680
Qy	478	GTGATCAGCAATCATGCGCAACATTTAGCAATCTTCTTATCCCAATGGGTTTCCAT	537
Db	681	GTAACCAACCAATGCATCGGCAACATTTAGCAACCTTTCGTATCCCAAGGATTTCCAT	740
Qy	538	TCTAAATACGAAAATTTTGTGTGATACAAATAAAAAATACTATATTATTTTTCAGATGCAAT	597
Db	741	TCTAATATGAAAACCTTTGTAGATAGCATAAAGATTAATATAAGTAACCGATCGAAT	800
Qy	598	GGATATTGGAATTTGAAAGGAGCAAAACCAGAAAGATTGGATTGAAGGAGCAGCGTAGCA	657
Db	801	GGGATTGGAACATGAAAGGATACGAATCCAGAAGATTGGATTATCATGAGCGGCAGTATT	860
Qy	658	GCTAAACAAGATTATCTTGGCGTGTGTAACGATACGACAAAGATTGCTTTGTAAGAACA	717
Db	861	GCGAAACAAGATTACGCTGGCATTTGTAATGATATAATACGAAAGATTGGTTCGTAAGACA	920

Qy	718	CCGCTATCTCAAGAAATATGAGATAAATGGCGTGGGAAGTAACACCCGGTGACAGGAAG	777
Db	921	GCTGTTTACCAAGAATATGAGATAAATGGCGCGCTGAAGTTACACCAATGACAGGTAAG	980
Qy	778	CGTTTAATGGAAGCGAGCGGTTACAGCTGCTTATATTTTGTGTTTGATACGAT	837
Db	981	CGTTTAATGGAATGACACACGTTTACTGCTGATACATTACAGTTTGGTTTGTATACGAT	1040
Qy	838	GTAAATCGCTAA	849
Db	1041	GGAATCGTTAA	1052
RESULT 3			
BCGENE			
LOCUS	BCGENE	2349 bp	DNA linear BCT 26-JUN-1997
DEFINITION	B.cereus genes plc and sph.		
ACCESSION	X64141		
VERSION	X64141.1	GI:312905	
KEYWORDS	phospholipase C; PLC gene; sph gene; sphingomyelinase.		
SOURCE	Bacillus cereus		
ORGANISM	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.		
REFERENCE	1 (bases 1 to 2349)		
AUTHORS	Gavrilenko, I.V., Baida, G.E., Karpov, A.V. and Kuz'min, N.P.		
TITLE	Nucleotide sequence of phospholipase C and sphingomyelinase genes from Bacillus cereus BKM-B164		
JOURNAL	Bioorg. Khim. 19 (1), 133-138 (1993)		
MEDLINE	93249510		
PUBMED	8387306		
REFERENCE	2 (bases 1 to 2349)		
AUTHORS	Kuzmin, N.P.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-JAN-1992) N.P. Kuzmin, Inet. of Biochemistry & Physiology of Microorganisms AS USSR, Pustchino, MOSCOW Region, 142292, USSR		
COMMENT	See also x64140.		
FEATURES	Location/Qualifiers		
source	1..2349		
	/organism="Bacillus cereus"		
	/mol_type="genomic DNA"		
	/strain="VKM 164"		
	/db_xref="taxon:1396"		
	/clone="pHR3"		
	41..45		
RBS	53..904		
gene	/gene="plc"		
CDS	53..904		
	/gene="plc"		
	/EC_number="3.1.4.3"		
	/codon_start=1		
	/evidence=experimental		
	/transl_table=11		
	/product="phospholipase c"		
	/protein_id="CAA45502.1"		
	/db_xref="GI:312906"		
	/db_xref="GOA:P09598"		
	/db_xref="SWISS-PROT:P09598"		
	/translation="MKKVLAAAIITVVAPLQVAFAHENDGSKIKIVHWSAEDK HKEGVNSHLWVNRADIMSRNTTLVKQDRVALLNEMWTELENGIYAADYENPYDYS TFASHFDPNKGTIPIYAKQAKETGAKYFKLAGESYKNDKMQAFYFLGSLHYLGD VNQPMHAANFNLSYPOGFHSHKYENFVDTIKDNKYVDGNGYWNWKGTPEDWIHGAA VVAQDYAGI VINDNTKDFVRAAVSQEVADKWRAEVTPMTCKRLMDAQRVTAGYIQLW FDTVGNR"		
	965..968		
RBS	/evidence=experimental		
gene	978..1979		
	/gene="sph"		
CDS	978..1979		
	/gene="sph"		
	/EC_number="3.1.4.12"		
	/codon_start=1		

Db	833	CGATTAAATGGATGCACAAAGCTGTTACTGCTGGATACATTCAGCTTTGGTTGTATACGTAC	892
QY	838	GTAAATGCGCTAA	849
Db	893	GGAGATCGTTAA	904
RESULT 4			
BCPLCSMC	BCPLCSMC	2700 bp	DNA linear BCT 26-JUN-1997
LOCUS	B. cereus genes for phospholipase C and sphingomyelinase C.		
DEFINITION	X12854 M20194		
ACCESSION	X12854.1 GI:39444		
VERSION	phospholipase C; PLC gene; smc gene; sphingomyelinase.		
KEYWORDS	Bacillus cereus		
SOURCE	Bacillus cereus		
ORGANISM	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.		
REFERENCE	1 (bases 1 to 2050)		
AUTHORS	Johansen,T., Holm,T., Guddal,P.H., Sletten,K., Haugli,F.B. and Little,C.		
TITLE	Cloning and sequencing of the gene encoding the phosphatidylcholine-preferring phospholipase C of Bacillus cereus		
JOURNAL	Gene 65 (2), 293-304 (1988)		
MEDLINE	88313678		
PUBMED	3137122		
REFERENCE	2 (bases 1441 to 2700)		
AUTHORS	Johansen,T., Haugli,F.B., Ikezawa,H. and Little,C.		
TITLE	Bacillus cereus strain SE-1; nucleotide sequence of the sphingomyelinase C gene		
JOURNAL	Nucleic Acids Res. 16 (21), 10370 (1988)		
MEDLINE	89057484		
PUBMED	2848222		
REFERENCE	3 (bases 1 to 2700)		
AUTHORS	Johansen,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-SEP-1988) Johansen T., Institute of Medical Biology, Cell Biology, University of Tromsø, 9001 Tromsø, Norway		
COMMENT	On Jul 24, 2002 this sequence version replaced gi:143336.		
FEATURES	Data kindly reviewed (11-JAN-1989) by Johansen T.		
source	Location/Qualifiers		
misc_feature	1..2700	/organism="Bacillus cereus"	
		/mol_type="genomic DNA"	
		/strain="SE-1"	
		/db_xref="taxon:1396"	
terminator	1..305	/note="ORF1"	
	291..340	/note="put. rho-independent terminator"	
misc_feature	368..445	/note="intergenic palindromic sequence"	
	453..458		
	476..481		
	545..554		
-35_signal		/note="pot. rRNA binding site"	
	562..1413		
-10_signal		/codon_start=1	
		/transl_table=11	
RBS		/product="pLc"	
		/protein_id="CAA31332.1"	
CDS		/db_xref="GI:295765"	
		/db_xref="GOA:P09598"	
		/db_xref="SWISS-PROT:P09598"	
		/translation="MKKVLALAAITVAPLQSVAFAHENDGGSKI KIVHRWSAEDK HKEGVNSHLWTVNRAIDIMSNTLVQDRVLAQNLNWEKTELENGIYAADYENPYDNS TFASHPYDNGKTVIPEAKOKETGAKYFKLAGESYKNKDMKQAFYVLGSLHYLGD VNPQMHAAFNLSYPQGFHSKYENFVDTIKDNYKVTGNGYNNWKGTNPWEIHGAA VVAQDYSIGVINDNTKQWFWKAAVSQETADKWRABVTMTGKRLMDAQRVTAGYQLW FDYTGDR"	
sig_peptide	562..633	/note="put. signal peptide (pLc)"	
	676..1410		
mat_peptide			

/evidence=experimental			
/transl_table=11			
/product="sphingomyelin phosphodiesterase"			
/protein_id="CAA45503.1"			
/db_xref="GI:580755"			
/db_xref="GOA:P09599"			
/db_xref="SWISS-PROT:P09599"			
/translation="MKGKLLGVLPFGIGLVGLVGGSSVQADTSTDQNTLKUMTHNV YMLSTNLYPNMQSRADLIGADYIKNQDVILNEVFDNSADRLIGNLKKEYPNQT AVLGRSNGNDKTLGYSSTPREDGVAIVSWKPIVEIKIYVFAKCGPDNLSNKG VYTKIKNDREPVHVIQTHLQAEQSMCKTSPASVRTNQLKEIQDFIKNQIPIINDEYVL PGDMYVKNIAENNSDEYASMEKTLHASIPSTGHTATWDATNTNIAKYNFSDSA EYLDIYIAKSDHANPSFIENKVLQPKSPQWTVTSWLKYYTYDDYSDHVPVAATISMK"			
ORIGIN			
Query Match	65.5%;	Score 556;	DB 1; Length 2349;
Best Local Similarity	79.1%;	Pred. No. 1.1e-104;	
Matches	674;	Conservative	0; Mismatches 175; Indels 3; Gaps 1;
1	ATGAAAAGAAAGTATTAGCACTAGCAGCTATGTTGCTTTAGCTGCGCCAGTTCAAAGT	60	
53	ATGAAAAGAAAGTACTTGTCTTTAGCAGCAGCTATTACAGTAGTAGCTCTTTTACAAAGC	112	
61	GTAGTATTGGCAACAAATAATAGTGAAGTCTCGCACCGGATTTTAA---GATGGTCA	117	
113	GTTCATTGCTCATGANAATGATGGGAGTAAATAAATAATAGTTCACCGCTGGTCT	172	
118	GCTGAGGATAGCATATAGAGGGATTAACCTCTCATTTGTGGATTGTAATCGTCAATT	177	
173	GCTGAAGATAAACAATAAGAGGTGTAATCTCATTTATGATGTTGTAACCGTCCGATT	232	
178	GACATCATGTCGTGAATACACAGATTGTGAATCCGAATGAACCTGCATTATTAATCAG	237	
233	GATATTATGTCCTCGCAATACACACTTGTGTAACCAAGATCGAGTTGCACAATTAATGAA	292	
238	TGGCGTGTGATTTAGAAAATGGTATTATTCTGCTGATTACGAGATCCTTATTATGAT	297	
293	TGGCGTACGAGTTAGAACCGTATTATTGCTGCTGCTACTAGAAATCCTTATTATGAT	352	
298	AATAGTACATATGCTTCTCATTTTATGATCCGGATACCTGGAACAACATATATCTCTTT	357	
353	AATAGCACATTTGCTTCACTTCTATGATCCAGACAATGGAAAAACATATATTCATTT	412	
358	GCGAAACATGCAAAAGAACGCGCAAAATATTTTAACTTGTGCTCAAGCATACCAA	417	
413	GCAAGACGCGCAAAAGAACTGGCGCTAAATATTTTAAATTAGCTGCTGATCATACAA	472	
418	AATCAAGATATGCGCAAGCATCTTCTACTTAGGATTATCGTTCAATTTATTAGGAT	477	
473	ATAAAGATATGAACAAGCATCTTCTATTAGGATTATCTCTTCAATTTATTAGGAT	532	
478	GTGAATCAGCAATGCAATGCAAGCAACTTTACGAATCTTCTTATCCATGGGTTCCAT	537	
533	GTAAACCAACCGATGATCGCGCAACTTTTACAAATCTTTTCATATCCCAAGGATTCAT	592	
538	TCTAATACGAAATTTGTTGTATACATAAATAAATAAATACTATATTGTTTCAGATAGCAT	597	
593	TCTAATATGAAAACCTTTGTAGATACGATAAAGATATTTAAGTAACGGATGGAAT	652	
598	GGATATTGGAATTGGAAGAGCAAAACCCAGAAGATTGGATTGAAGAGCAGCGGTAGCA	657	
653	GGATATTGGAATCGGAAGAGTACAAATCCAGAAGAGTGGATTATCGAGCGCGAGTAGTA	712	
658	GCTAAACAAGATTATCTCGCGGTGTGAAACGATACGACAAAGATTGGTTGTAAAGCA	717	
713	GCGAAACAAGATTACTCTCGGAATTTGTAATGATTAATACGAAGATTGGTTGCTGAAGCA	772	
718	GCGGTATCTCAAGATATGCAAGATAAATGGCGTGGCAAGTAAACACCGGTGACAGAAAG	777	
773	GCTGTGTACAGAATAATGACAGATAAATGGCGCGCTGAAAGTTACCAACCAATGACAGGTAAG	832	
778	CGTTTAAATGGAAGCGCAGCGGTTACAGCTGGTTATATTCATTTGTTGGTTGATACGAT	837	

```
terminator      /product="PLC"
1413..1465      /note="mature PLC exoenzyme"
1418..1423      /note="pot. rho-independent terminator"
1441..1446      /note="pot. rho-independent terminator"
1471..1479      /note="pot. rRNA binding site"
1487..2488      /codon_start=1
/product="SMase"
/protein_id="CAA3133.1"
/db_xref="GI:580761"
/db_xref="GOA:P09599"
/translation="MKGKLLKGLVSGFGLVGLVGGSSVQADSTDDNNTLKVMTHNV
YMLSTNLYPNQGSORADLIGAADYIKNQDVILNEVDNSADRLGLNKLKYPNOT
AVLGRSNGENDKTLGVSYSSTPEDGGVAIVSKPIVEKIQYVFAKGGCPNLSNKG
VYTKKGNDRFVHIGTHLQADSMCGKTSKSPASVRLNOLKEIQDFKKNKIPNDYVL
FGDMNVNKNIAENNSDEYASMFKTLHSLPSYTGHTATWDATNLSAKYFPDSPA
EYLDYITASKDHANPSPFIENKLPQSPQWTVSLKKYTYDDYSDHYPAATISMK"
1487..1567      /note="put. signal peptide (SMase)"
1568..2485      /product="SMase"
/note="mature SMase exoenzyme"
2493..2528      /note="pot. rho-independent terminator"

ORIGIN
Query Match      65.5%; Score 556; DB 1; Length 2700;
Best Local Similarity 79.1%; Pred. No. 1.1e-104;
Matches 674; Conservative 0; Mismatches 175; Indels 3; Gaps 1;

QY 1 ATGAAAAGAAAGATTAGCACTAGCAGCTATGTTGCTTTAGCTGCCAGTTCAGAGT 60
Db 1 ATGAAAAGAAAGATTAGCACTAGCAGCTATGTTGCTTTAGCTGCCAGTTCAGAGT 60
QY 562 ATGAAAAGAAAGATTAGCTTTAGCAGCAGCTATTACAGTAGTAGCTCCCTTACAAAGC 621
Db 562 ATGAAAAGAAAGATTAGCTTTAGCAGCAGCTATTACAGTAGTAGCTCCCTTACAAAGC 621
QY 61 GTAGTATTTCGCAAAACAAATAATAGTGAAGATCTCTGCACCGCATTTTAA--GATGGTCA 117
Db 61 GTAGTATTTCGCAAAACAAATAATAGTGAAGATCTCTGCACCGCATTTTAA--GATGGTCA 117
QY 622 GTTGCAATTGCTCATGAATAATGATGGGGAGTAAATAAATAATAGTTCCACCGCTGGTCT 681
Db 622 GTTGCAATTGCTCATGAATAATGATGGGGAGTAAATAAATAATAGTTCCACCGCTGGTCT 681
QY 118 GCTGAGGATAAGCAATAATGAGGGGATTAATCTCTCAATTTGTGATTTGTAATCGTCAAT 177
Db 118 GCTGAGGATAAGCAATAATGAGGGGATTAATCTCTCAATTTGTGATTTGTAATCGTCAAT 177
QY 682 GCTGAAGATAAACAATAAAGAGGTGTAATCTCATTTATGATGTTGTAACCGTCCGATT 741
Db 682 GCTGAAGATAAACAATAAAGAGGTGTAATCTCATTTATGATGTTGTAACCGTCCGATT 741
QY 178 GACATCATGCTCGTAATACACGATTTGATCCGAATGAACTGCAATTAATGATGAG 237
Db 178 GACATCATGCTCGTAATACACGATTTGATCCGAATGAACTGCAATTAATGATGAG 237
QY 742 GATATTATGCTCGCAATACACACTTGTAAACCAAGATCGAGTTGCACAATTAATGAA 801
Db 742 GATATTATGCTCGCAATACACACTTGTAAACCAAGATCGAGTTGCACAATTAATGAA 801
QY 238 TGGCTGCTGATTAGAAAATGATTTATTTCTGCTGATTAGCAGAACTCTTATTATGAT 297
Db 238 TGGCTGCTGATTAGAAAATGATTTATTTCTGCTGATTAGCAGAACTCTTATTATGAT 297
QY 802 TGGCTGCTGATTAGAAAATGATTTATTTCTGCTGATTAGCAGAACTCTTATTATGAT 861
Db 802 TGGCTGCTGATTAGAAAATGATTTATTTCTGCTGATTAGCAGAACTCTTATTATGAT 861
QY 298 AATAGTACATATGCTTCTCATCTTTATGATCGGATACCTGGAACCAATATATCTCTTT 357
Db 298 AATAGTACATATGCTTCTCATCTTTATGATCGGATACCTGGAACCAATATATCTCTTT 357
QY 862 AATAGCATTGCTTCTCATCTTTATGATCGGATACCTGGAACCAATATATCTCTTT 921
Db 862 AATAGCATTGCTTCTCATCTTTATGATCGGATACCTGGAACCAATATATCTCTTT 921
QY 358 GCGAACAATGAAAGAACAGGGCGCAAAATATTTTAACTTGTGCTGCTGCAACATACCA 417
Db 358 GCGAACAATGAAAGAACAGGGCGCAAAATATTTTAACTTGTGCTGCTGCAACATACCA 417
QY 922 GCAAAGCAGGCAAAAGAACTGGCGCTAAATATTTTAAATGATGCTGGTGAATCATACAA 981
Db 922 GCAAAGCAGGCAAAAGAACTGGCGCTAAATATTTTAAATGATGCTGGTGAATCATACAA 981
QY 418 AATCAAGATATGCAAGCAAGCTTCTTCTACTTAGGATTTAGCTTCAATTTATTAGGAT 477
Db 418 AATCAAGATATGCAAGCAAGCTTCTTCTACTTAGGATTTAGCTTCAATTTATTAGGAT 477
QY 982 AATAAAGATATGAACAAGCAAGCTTCTTCTATTAGGATTTATCTTCAATTTATTAGGAT 1041
Db 982 AATAAAGATATGAACAAGCAAGCTTCTTCTATTAGGATTTATCTTCAATTTATTAGGAT 1041
QY 478 GTGAATCAGCCCAATCATGCGCAAACTTTACGAACTCTTCTTATPCCAATGGGTTTCCAT 537
Db 478 GTGAATCAGCCCAATCATGCGCAAACTTTACGAACTCTTCTTATPCCAATGGGTTTCCAT 537
QY 1042 GTAAACCAACCGATGATGCGGCAAACTTTACAAATCTTTCATATCCCAAGGATTCCAT 1101
Db 1042 GTAAACCAACCGATGATGCGGCAAACTTTACAAATCTTTCATATCCCAAGGATTCCAT 1101
QY 538 TCTAATACGAAATTTTGTGATACATAAATAAATACTATATATTTGTTTCAGATGCAAT 597
Db 538 TCTAATACGAAATTTTGTGATACATAAATAAATACTATATATTTGTTTCAGATGCAAT 597
```

```
Db 1102 TCTAATATGAAACTTTTGTAGATACGATAAAGATAATATTAAAGTACCGATGGAAT 1161
QY 598 GGATATTGGAATTTGAAAGAGGACAAACCCAGAGAATTTGGATTGAAAGGACGCGTAGCA 657
Db 1162 GGATATTGGAATTTGAAAGAGGACAAACCCAGAGAATTTGGATTGAAAGGACGCGTAGCA 1221
QY 658 GCTAAACCAAGATTTATCTGCGGTTGTGAACATACGACAAAAGATTGTTGTTAAAGCA 717
Db 1222 GCGAAACCAAGATTTATCTGCGGTTGTGAACATACGACAAAAGATTGTTGTTAAAGCA 1281
QY 718 GCGGTATCTCAAGATATGTCAGATAAATGGCGTGGGAAGTAAACACCGGTGACAGGAAAG 777
Db 1282 GCTGTGTCACAGAATATGTCAGATAAATGGCGCTGAAGTTACACCAATGACAGGTAG 1341
QY 778 CGTTTAATGGAAGCGCAGCGCTTACAGCTGGTTATATTTGTTGTTGATACGTAT 837
Db 1342 CGATTAAATGGAATGTCAGACAGTGTACTGCTGATACATTCAGCTTTGTTGATACGTAT 1401
QY 838 GTAATATCGCTAA 849
Db 1402 GGAGATCGTTAA 1413

RESULT 5
BCFUSE      BCFUSE      2201 bp      DNA      linear      BCT 17-JUN-1993
LOCUS      B.cereus fuse genes plc and sph (=h-cerAB)
DEFINITION      X64140
ACCESSION      X64140
VERSION      X64140.1 GI:39421
KEYWORDS      cereolysin AB, fuse genes.
SOURCE      Bacillus cereus
ORGANISM      Bacillus cereus
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE      1
AUTHORS      Gavrilenko, I.V., Baida, G.E., Karpov, A.V. and Kuz'min, N.P.
TITLE      Nucleotide sequence of phospholipase C and sphingomyelinase genes from Bacillus cereus BKM-B164
JOURNAL      Bioorg. Khim. 19 (1), 133-138 (1993)
MEDLINE      93249510
PUBMED      8387306
REFERENCE      2 (bases 1 to 2201)
AUTHORS      Kuzmin, N.P.
TITLE      Direct Submission
JOURNAL      Submitted (14-JAN-1992) N.P. Kuzmin, Inst. of Biochemistry & Physiology of Microorganisms AS USSR, Pustchino, Moscow Region, 142292, USSR
COMMENT      See also X64141.
FEATURES      location/Qualifiers
source      1..2201
/organism="Bacillus cereus"
/mol_type="genomic DNA"
/strain="VKM 164"
/db_xref="taxon:1396"
/clone="pF27p"
41..45
/note="SD h-cerAB"
53..1831
/gene="h-cerAB"
53..1831
/gene="h-cerAB"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="hybrid cereolysin AB"
/protein_id="CAA45501.1"
/db_xref="GI:39422"
/db_xref="GOA:P09598"
/db_xref="GOA:P09599"
/db_xref="SWISS-PROT:P09598"
/db_xref="SWISS-PROT:P09599"
/translation="MKKKVLAALAAITVAPLQSVAFAHENDGSKIKIVRWSAEDK
HKEGVNSHLWTVNRAIDMSRNTTLVKQDRVAQLNWRLENGIYAADYENPYDINS
```

TFASHFVDPNGKTYIIPFAQAKETGAKYFKLAGESYKNKDMKQAFVYLGSLHLYGD	
VNQPMHAANFTNLSYPOGFHSKYENFVDITIKONTKVTDGNGYWNWKTNEERWIHGA	
VVAQDYSGIVNDNTKQWFKVAQSVQBYADKRAEVTPTMTGKRLMDAQRVTAIGYIQW	
FDYTGDRSSRVDPDQNNTLKVMTHYMLSTNLYPNWQSQRADLIGAADYIKNOOV	
VILNEVPDNASDRLLGNLKEYPNQTAVLGRSNGNEWKTLGYSSTSPEDGGVAIV	
SKWPIVEKIOVPAKCGPDLNLSNKGFTTKIKKNDRFVHVIGTHLQAEBSMCKTSP	
ASVTNQLKEIQDFIKNNIPNDEYVLPFGDMVYNKINAENSDSEYASPKTLHASI	
PSYTGHTATWDATNSIAKYNFPDSPAIEYLDIYIIASKDHANPSEIENKVLQPKSPQT	
VTSWLKXYTYDDYSDHPVAATISMK	
53. .899	
misc_feature	/gene="h-cerAB"
	/note="structural part of gene plc"
	/evidence=experimental
900. .917	
misc_feature	/gene="h-cerAB"
	/note="linker"
	/evidence=experimental
918. .1828	
misc_feature	/gene="h-cerAB"
	/note="structural part of gene sph"
	/evidence=experimental
ORIGIN	
Query Match 65.3%; Score 554; DB 1; Length 2201;	
Best Local Similarity 79.1%; Pred. No. 2.8e-104;	
Matches 672; Conservative 0; Mismatches 175; Indels 3; Gaps 1;	
QY	1 ATGAAAAAGAAAGTATTAGCACTAGCAGCATGTTGGTCTTTAGTCGCCCAAGTTCAAAGT 60
DB	
QY	53 ATGAAAAAGAAAGTACTTCTTTAGCAGCAGCATATTACAGTAGTAGCTCTCTTTACAAAGC 112
DB	
QY	61 GTAGTATTGGCAACAACAAATAATAGTAGTCTCGCACCGATTTTAA---GATGGTCA 117
DB	
QY	113 GTTGCAATTGGCTCATGAAATAATGATGGGGAAGTAAATAAATAATAGTTCACCGCTGGTCT 172
DB	
QY	118 GCTGAGGATAAGCATAAATGAGGGGATTAATCTCATTTGTGATTTGTAATCGTGCATTT 177
DB	
QY	173 GCTGAAGATAAATGAGAGGTGTAAATCTCATTTATGATTTGTAACCGTGGATTT 232
DB	
QY	178 GACATCATGTCCTGTAATCAACAGATTGTGAATCCGAATGAACTGCAATTATTAAATGAG 237
DB	
QY	233 GATATATTGCTCGCAATACACACTTGTAAACCAAGATCGAGTTGCAAAATTAATGAA 292
DB	
QY	238 TGGCTGCTGATTTAGAAATGGTATTATTCTGCTGATTCAGAACTCTTATTATGAT 297
DB	
QY	293 TGGCGTACGAGTTAGAACCGGTATTATTGCTGCTGACTATGAAATCTTATTATGAT 352
DB	
QY	298 AATAGTACATATGCTTCTCACATTTTATGATCCGGATCTGGAACACATATATTCCTTTT 357
DB	
QY	353 AATAGCACATTTGCTTCACATTTCTATGATCCAGACATGGAAAAACATATATTCATTT 412
DB	
QY	358 GCGAAACATGCAAAAAGAAACAGCGCGCAATATTTTAACTTCTGCTGGTCAAGCATACCA 417
DB	
QY	413 GCAAAAGCAGCAAAAAGAACTGGCGCTAAATATTTTAAATAGCTGGTGAATCAACAAA 472
DB	
QY	418 AATCAAGATATGAGCAAGCATTTCTTCTACTTAGGATATCGCTCATATTATTAGAGAT 477
DB	
QY	473 AATAAAGATATGAAACAGCATCTTCTATTAGGATATCTCTTCATATTATTAGAGAT 532
DB	
QY	478 GTGAATCAGCCCAATGATCGACAAACTTTTACGAATCTTTCTTATCCAAATGGGTTTCCAT 537
DB	
QY	533 GTAAACCAACCGATCATCGGCAAACTTTTACAAATCTTTTATATCCACAGGATTTCCAT 592
DB	
QY	538 TCTAAATAGCAAAATTTTGTGTATACAAATAAATAAATACTATATTGTTTTCAGATAGCAAT 597
DB	
QY	593 TCTAAATATGAAATCTTTGTAGATACGATAAAAGATAATATAAAGTAACCGATGGAAT 652
DB	
QY	598 GGATATTGGAATTGGAAGAGCAACCCAGAGATTGGATTGAAGAGCAGCGGTAGCA 657
DB	
QY	653 GGATATTGGAATCTGGAAGAGTACAAATCCAGAGAGTGGATTTCATGAGCGGCAGTAGTA 712
DB	
QY	658 GCTAAACAGATTTATCTCTGCGCTTTGTGAACGATACACAAAGATTTGGTTTGTAAAGCA 717
DB	

Db	713	GCAGAAACAAGATTACTCTCGAATTGTAAATGATAATAACGAAAGATTGGTTTCGTAAGAACA	772
QY	718	GCCTATCTCAAGAATATCGATTAATCGCTGCGGAGTAGTACACCGGTGACAGGAAAG	777
Db	773	GCTGTGTCAACAAGAAATGTCAGATAAATGCGCGCTGGAAGTTTACCAATGACAGGTAAG	832
QY	778	CGTTTAAATGAACGCGAGCGGCTTACAGCTGGTTATATTCATTGTTGGTTTGTATACGTAT	837
Db	833	CGATTAAATGATGACACACGTTACTGCTGGATACATTACGCTTTGTTTGTATACGTAC	892
QY	838	GTAAATCGCT 847	
Db	893	GGAGATCGAT 902	
RESULT 6			
AE017026		290214 bp	DNA linear BCT 30-APR-2003
LOCUS		Bacillus anthracis str. Ames section 3 of 18 of the complete	
DEFINITION		genome.	
ACCESSION	AE017026	AE016879	
VERSION	AE017026.1	GI:30254178	
KEYWORDS		Bacillus anthracis str. Ames	
SOURCE		Bacillus anthracis str. Ames	
ORGANISM		Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus; Bacillus cereus group.	
REFERENCE		1 (bases 1 to 290214)	
AUTHORS		Read, T., Peterson, S., Tourasse, N., Baillie, L., Paulsen, I., Nelson, K., Tettelin, H., Fouts, D., Eisen, J., Gill, S., Holtzapple, E., Okstad, O., Helgason, E., Rillstone, J., Wu, M., Kolonay, J., Beanan, M., Dodson, R., Brinkac, L., Gwinn, M., DeBoy, R., Madupu, R., Daugherty, S., Durkin, A., Haft, D., Nelson, W., Peterson, J., Pop, M., Khouri, H., Radu, D., Benton, J., Mahamoud, Y., Jiang, L., Hance, I., Weidman, J., Berry, K., Plaut, R., Wolf, A., Watkins, K., Nierman, W., Hazen, A., Cline, R., Redmond, C., Thwaite, J., White, O., Salzberg, S., Thomas, B., Friedlander, A., Koehler, T., Hanna, P., Kolsto, A.-B. and Fraser, C.	
TITLE		The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria	
JOURNAL		Nature 423 (5935), 81-86 (2003)	
MEDLINE		22608414	
PUBMED		12721629	
REFERENCE		2 (bases 1 to 290214)	
AUTHORS		Read, T., Peterson, S., Tourasse, N., Baillie, L., Paulsen, I., Nelson, K., Tettelin, H., Fouts, D., Eisen, J., Gill, S., Holtzapple, E., Okstad, O., Helgason, E., Rillstone, J., Wu, M., Kolonay, J., Beanan, M., Dodson, R., Brinkac, L., Gwinn, M., DeBoy, R., Madupu, R., Daugherty, S., Durkin, A., Haft, D., Nelson, W., Peterson, J., Pop, M., Khouri, H., Radu, D., Benton, J., Mahamoud, Y., Jiang, L., Hance, I., Weidman, J., Berry, K., Plaut, R., Wolf, A., Watkins, K., Nierman, W., Hazen, A., Cline, R., Redmond, C., Thwaite, J., White, O., Salzberg, S., Thomas, B., Friedlander, A., Koehler, T., Hanna, P., Kolsto, A.-B. and Fraser, C.	
TITLE		Direct Submission	
JOURNAL		Submitted (26-MAR-2003) The Institute for Genomic Research, 9712	
FEATURES		Medical Center Dr, Rockville, MD 20850, USA	
source		Location/Qualifiers	
		1. 290214	
		/organism="Bacillus anthracis str. Ames"	
		/mol_type="genomic DNA"	
		/strain="Ames"	
		/db_xref="taxon:198094"	
		140. .949	
		/locus_tag="BA0573"	
		140. .949	
		/locus_tag="BA0573"	
		/note="Identified by Glimmer2; putative"	
		/codon_start=1	
		/transl_table=11	
		/product="hypothetical protein"	
		/protein_id="AAP24593.1"	
		/db_xref="GI:30254179"	
gene			
CDS			

```

/translation="MMVFTIGVMSLSILGASPSSETSCKTKTDYKQSKQKQNNKGIH
AVHTEAKELGIEGTQKEQITLEKEIHETEVEGRAKQIGHSIECKQVDTLSEETIET
KVEQALKLIGSIENTSINLINOINTKINDEADKLGTNGKEIEDIAEIEYQTKV
REAGKLGISPKKEIEVLAQEVYEQKQEEAKEVHIDLYGKDIQYVLSIEINEQKVLQ
MADELNMKTNMNVQELAEKIKKQPEKGKELNFPVVRTDADAFYSYLTN"
/locus_tag="BA0574"
963..1709
/locus_tag="BA0574"
963..1709
/locus_tag="BA0574"
/translation="similar to GP:16410113, SP:P07114, GB:S96966,
GB:X04830, and PID:42533; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP24594.1"
/db_xref="GI:30254180"
/translation="MIRKFKVVICWALIGIAGVGIYNGITKYVEKLTFFNIQSGKOV
EFNRDEKVLNVNPIQQLPELDRCEVTSIAMLQYAGITVDKMKLANEIKKVDFFNM
DVGKRNPEGFVNIYTSSEGYGVYHGLFQLAKKYLPNKXAVDITGKSIKELYKSVK
AGQPVIITNATFAFLDEBFTWETNNGDVSTYNERHCVVLLIGYDQESVYTRDPLKD
SLDVKVPKEPEQAWQVMSQAISYVKRSK"
/locus_tag="BA0575"
1963..3945
/locus_tag="BA0575"
1963..3945
/locus_tag="BA0575"
/translation="similar to GB:Z14244, SP:P24311, PID:30151,
GB:Z14244, SP:P24311, and PID:30151; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="methyl-accepting chemotaxis protein"
/protein_id="AAP24595.1"
/db_xref="GI:30254181"
/translation="MQKIFNVKQMSIKKKLIISPLILITIPGLIGVSYQAKTNF
EQQITGKAKENISILNTVISQNIIEKFVDAAYFADILTEDTYPNGQKSVITKLAQYI
KLHPEVEGIYIGTGKFIREFPIQMSDGINPTDRDWTYKEAHENKGVIVTAPYQAS
TKNMVVTIAVKQGVIGINLNDNLKISMINIGERGVAVILDDQNKQIVSHPSR
KPGKAVTDPKVPIYEDQKRNVAYTEKDDKKNLIPATNEKTGWKVGVMFDEITQAA
NPVFKTIVIAIVFQSVLIYFTKSTIPLRKIADSAVKISKGLDTEKITIHSKD
DIGKLSFNMSASLODVIQISFASHVAAASAEELTASVQOANDADQITIAMQOV
SGKASQCVQEEGNATLQVNTALQVNGSASISISLHARERAKGEBLVEGTAT
QMQSISRSVSSDAILKLLDESKSQVGAISAIQNIATQTNLLALLAIAEIAARAGEQG
RQFAIVADSVKRAHQSSGSEGEIANLJAEI KADIEHTVKAMDNNVGVQGLDVVTK
TKVSTELSTHTHVSQVNMVETTKIAGDANEVINAIIDEIAAAABSENTASMQSVA
ASTBEQVMSMEIISASQSLAEAEELQAMTSKPKV"
/locus_tag="BA0576"
4024..5628
/locus_tag="BA0576"
4024..5628
/locus_tag="BA0576"
/translation="similar to GP:10173010; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="sensory box histidine kinase"
/protein_id="AAP24596.1"
/db_xref="GI:30254182"
/translation="MKKRLNWLWKTTITLVCTVVIPLSLVTDILISHNVERTTDS
QAEKAKTAHVANDSIVDGLIGKADTSAIQTYNRLKNTGVQVIVVMDMNGIRKS
HPNPKIHHFIFGGDEGALKGKHEVSLAEGTLGISMRFVVPFIFSETGEQLGAVAGI
SADNPKVKEKSRHIYIGVGVVLVIGIALLARHKKSLFGLPHRIAKIIEERN
TMLQSVKEGIITAVDKAEAVTLINNEAKRLFKKSGLEEDFIGKDVLEYPNERSIKEVLQ
TGEVOLNEQNIYGITIVTNVPLVKEIEICAIATFRDKTEIRKLAELTGIRLYAE
ALRAQSHBFMKMHVVLGLTHMKQVEELQKISGMVSEHVEIGVMKRIKSPVPFAGV
LLGKLSYAREKNKILISEDSYMPEIDDESITHELITIVGNLINDALEAVTNCKKRV
EVKIQHEDILITQDTGKGIOEKIEBELFTNGSKGDNKGNGYGLYLKVESIQRINGE
IHMHSLVKGKGTITITIEIPKGRDQI"
/locus_tag="BA0577"
5658..6332
/locus_tag="BA0577"
5658..6332
/locus_tag="BA0577"
/translation="similar to GP:1934810, SP:P26908, PID:40173, and
GB:AL009126; identified by sequence similarity; putative"

```

```

/codon_start=1
/transl_table=11
/product="response regulator"
/protein_id="AAP24597.1"
/db_xref="GI:30254183"
/translation="MVAMNLNTHYLEOVGFFELVQAVNSIKSAIEVLEESRIDLVLDLI
FMPETGPELLMYRNQKEKIDIMKISAVHDMGSIKKALQYGVVDYDLIKPTFFEPFKE
ALTIREKLTQKQCKQISQSELDLILQKREPTVTKELPKGTUKQTQLIWMQKIE
SLNGRVFTTDEMAQLVGISRSIRKYNMFLTDIGVLENNEMYQHVGREPVSKLRCVDQK
KIEFVV"
/locus_tag="BA0578"
6455..7801
/locus_tag="BA0578"
6455..7801
/locus_tag="BA0578"
/translation="similar to GP:1146122, SP:P15300, GB:X14135,
PID:42079, GB:U00096, PID:1651280, PID:1651286, and
PID:1786892; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="citrate cation symporter family"
/protein_id="AAP24598.1"
/db_xref="GI:30254184"
/translation="MGIQKNVEAVSFSENEVQRESFASKIMNVKIGVILPLPYVYVLA
AIIYCASVYNKLPADMIGCGFAVIMIGIFLGDIGHRIPILKNIIGPAILSLFIPSLLV
FFNMNPASMEAAATMDMKSNFLDIYISCLVGSILGNRRKVLVQGFVFMFIPLVVGT
LASVAVGLLVGSFGFEMKHTFFIIVIVSGGIGEGILPLSLAYSLDILNESSATFVS
QLIPAAIIGNFAIVSAGYMKRLGKPELPSGNGVLVKTQNAELKQNTKPKIDFS
LMGAGLIACTFFIFGGFASKFIPGAIIMFSAALVKYFKLMPAKMEQQGAFHLYKF
ISKNLTWPLMVGLLXIPLAKDVAALSVGVVACVSLVLTWASGFLVGVKMKVYPV
ESAIYVTSCHSLGSGTGDVAIILSASNRMLMPPAQISTRLGGANVVVTAIILLKMF5"
/locus_tag="BA0579"
7861..9060
/locus_tag="BA0579"
7861..9060
/translation="similar to GP:10173012, and GP:10173012;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="malate dehydrogenase, putative"
/protein_id="AAP24599.1"
/db_xref="GI:30254185"
/translation="MLANQINERSILLHKLGVLEITSKYEVNSADDLSLTYTPGVA
ESCKAIADEETVYDTARGNNVAVSDGTAVLGNIGIPKRAAMPVMEGKSLFKFKPA
NVDAPFLCQITVDEIVTLVKNLEPTFAGINLEDAAPRCFEIEKRLKEBTNIPVPH
DDQHTAIVLVAIVNALKVSKQNDHVKIVINGAGSAGIAGIKLLKAGAHITPLVS
LEGIVCEGAAMNEAQIEVSKTKNREHVRGTLEKAIHOADIPIGVSAENVLTKELVOM
MNEKAIYPMANPIPEIFPEDALEAGAAVCTGSDVANOVNNVLAAPGIRGALDVR
ATDIETEMKLAAYGIANIITDEENANNVIPNPLDKRVVFSVAEAAKAAIDSGVAG
IMKIPSY"
/complement(9356..9880)
/locus_tag="BA0580"
/complement(9356..9880)
/locus_tag="BA0580"
/codon_start=1
/transl_table=11
/product="lipoprotein, putative"
/protein_id="AAP24600.1"
/db_xref="GI:30254186"
/translation="MKKIGTMLLPSILLIAGCTAQAPDLKKPKKEAIAISTQVNPAPSF
PHLSVLKDVNNEETPSFQEKIPLKGEIEKIAMADSPLIANEKKNIMYFLDPENPTG
KLSIALQKQSVTPTPLLFQOESSEPTTTSITDSTTNTPLTMSLPSSGLWLNLIY
VNEKYDQFVITAE"
/complement(9996..10922)
/locus_tag="BA0581"
/complement(9996..10922)
/locus_tag="BA0581"
/translation="identified by match to PFAM protein family HMM
PF04173"
/codon_start=1
/transl_table=11
/product="transporter, BamA family"

```

Query Match

65.1%; Score 552.8; DB 1; Length 290214;


```
gene
CDS
/locus_tag="BC0619"
/codon_start=1
/transl_table=11
/product="Iron(III) dicitrate transport ATP-binding
protein fecB"
/protein_id="AAP07636.1"
/db_xref="GI:29894346"
/translacion="MKALETNRLTSLYGETIIDELNLEIPKGEITIPISNGCGKS
TLLSLARLKTPTGIDLLDNOAQISMQTKQIAERMAILPQGPQPEGLTVLQVKQG
RYPVTLKQSEKDEENVQKALATGTEFAERDVHALSGQRQRAWIAMTLAQDTG
IILLDEPTTLYDMTHQIEVLDDLPELNETEORTIWMVLDLNLACRAVDNIVAIQDKQ
IYAQCKPREVDEKLVRDVERMECQISTDPLFGTGLCPHGRGRVVKVEQAQMR"
complement(4285..5016)
/locus_tag="BC0620"
complement(4285..5016)
/locus_tag="BC0620"
/EC_number="2.1.1.-"
/codon_start=1
/transl_table=11
/product="Ubiquinone/menaquinone biosynthesis
methyltransferase UBis"
/protein_id="AAP07637.1"
/db_xref="GI:29894347"
/translacion="MAFTESDVYNTAFEOVMKRYRENSPESIEKPVFOLIGHV
QKQILDLGCGDAFGASLLEKGCYSYTGIGSELMYEKAKKQLENKNGSVHFLNLKD
YTFESSTDLVTSRLALHYIEHLDTFIQNVFTLKTNGTFTFSVQHVPVITSFSLQT
SGKRTIWDVDFKLGKRVPEWIDQEVIKYHRTTBEXFTLLQAGFTITSLKEATPNP
TYFQDEEYKRLRIPLFLFSCTK"
5230..6420
/locus_tag="BC0621"
5230..6420
/locus_tag="BC0621"
/EC_number="2.3.1.29"
/codon_start=1
/transl_table=11
/product="2-amino-3-ketobutyrate coenzyme A ligase"
/protein_id="AAP07638.1"
/db_xref="GI:29894348"
/translacion="MSSKTLAKLFLENLEDLSKGLYNNVIDPLESSNGPIITIGSKY
INLSNNYLGLTDSRLQEAAGAIHKGVGAGAVRTINGTLDLHLKLEETIAKPKHT
EAAIAYQSGFNCNMAAISAVMDKDAIISDELNHASIIDGSRLSKAKIIVVYKPMED
LRQALAKESGLYNKLMVITDGVFSDMGDVAKLPEIVEIAELDLMTYVDAGSGV
LKGAGTVKHFGLSDQFQIGTSLKALGVIGGVAGKQNLIDMLKVRSPREPLTGLD
TPADAACMRSEIEMSETELDHRLWENGRYLKQGLKELGNIGNESETPTIPFCIIGDE
VLTQFESRLNEEGYAKSIIVFPTVAKGTGRVRNPTAAHTKEMLDEAILKYEKVGKE
MGI"
6465..7430
/locus_tag="BC0622"
6465..7430
/locus_tag="BC0622"
/EC_number="1.1.1.103"
/codon_start=1
/transl_table=11
/product="L-threonine 3-dehydrogenase"
/protein_id="AAP07639.1"
/db_xref="GI:29894349"
/translacion="MKKILVTGSLQIGSELVLMKLRDVGASNVLIATDIRETDSVVT
SGPFETLDVQOKLIDIAKNEVDTIHLAALLSATAEKNPLPAWNLNMGVLNVALE
AARELNCKFFTPSSGAGPSTPKDNTQDPTQRPFTMYGVNKVAGELLCDYHQKFG
VDRGRVRFGLISYVAPGGGTTDYAVEIYEAKKGYATSYIAEGTMDMMYMPDAL
QAIISLEADPSKLVRHAFNITAMSFEPEQIAASIRKHIPTFTNDYAVDFPARQTIAD
SWPNSIDATAAMKENGPKAEYDLDKMTTMLAKLKKLTAELVMN"
7490..7912
/locus_tag="BC0623"
7490..7912
/locus_tag="BC0623"
/codon_start=1
/transl_table=11
/product="Phosphohydrolase (MutT/nudix family protein)"
/protein_id="AAP07640.1"
/db_xref="GI:29894350"
/translacion="MQRVDVVYALIHDEETKILMVHNVEQNWSLPGGAVEKGEILE
BALIREVKEETGLTATLGLVAINEKEFFESGNHALFTTFRANVEKGEILAEDEGEIS
```

```
gene
CDS
/locus_tag="BC0624"
/codon_start=1
/transl_table=11
/product="von Willebrand factor type A domain protein"
/protein_id="AAP07641.1"
/db_xref="GI:29894351"
/translacion="MROIQFSDKKIDASLFLOMENLMYALLKEDDAYLBYGVKAYDEI
EKKVISHFWDKKEEDTVIGLSEVFLKALGNKHYSMTLIRSVAILQSPKLPKFL
TQLFVLEDLRAVEIIVKMLRPGTKRHKRMYRNYFSSQNEINRVNYGDRFLFCL
CYLSUTSDKYEMFNNEPEQIEAILHETFOARNTEDIIMYIVKRIYRLEELLESMLN
NYFGHAPLYLSLADENCRVEDLANDDTQIIDDENKMDSEFSTWHRENKNNE
NLFRLFESGTTKNMGSTARESDGOALGSVOQTSKQSNFPGDADTEARASHA
SSONDGAYKVNTGASHTFEARPNDEKDYQAIKSVNVNKKVKELAKIITEKTENK
TNSDKYVGLRKKFLRIYTEKOPRMYKKGSOELDVAQLLVDCSSGMYNKMEI
TKSVLPLHEALKSLKIPHAISGFWEDASSAKPEDKPNVIEHVNYKNSTLPVGPPI
MQLREEDNRDGYIIRIVSEKLAKRPEKHFLLVTDGEPALDYOQDQILDTHBAVK
LARKSGMEVIGIFIEEGEAKBATYQLMKNYIHHFPLVANHAEDLRLKLPKLLKLLK
TIE"
complement(9837..10730)
/locus_tag="BC0625"
complement(9837..10730)
/locus_tag="BC0625"
/codon_start=1
/transl_table=11
/product="NorQ protein"
/protein_id="AAP07642.1"
/db_xref="GI:29894352"
/translacion="MLSFQHIADLKQOLSTHKKNOQAGENAHLIQTKIYKASDNS
IIEDAITALLGKNIILKPGTSGKTVLAETLSLPQKPMHSINCSVDLVEGILGN
TLOTKDGASESVFVNGPLKAMKNGHFLYIDEINMAKPETLPLHLGALDYRQWITNPF
TOEVVYGDVEYRVIAINEGVGTSELNEALKNRFVIEVPIQGDITLKEILLASLKL
NDVATIEKFVAFASDLPLARDGRVSEASIRGIIDACDLGVITPWRATERSIAK
LNDETQMTVRELAEISFFEG"
complement(10858..12387)
/locus_tag="BC0626"
complement(10858..12387)
/locus_tag="BC0626"
/EC_number="2.7.8.-"
/codon_start=1
/transl_table=11
/product="Cardiolipin synthetase"
/protein_id="AAP07643.1"
/db_xref="GI:29894353"
Query Match 65.1%; Score 552.8; DB 1; Length 301289;
Best Local Similarity 78.9%; Pred. No. 3e-104;
Matches 672; Conservative 0; Mismatches 177; Indels 3; Gaps 1;
Qy 1 ATGAAAAGAAAGTATTAGCAGCTAGCAGCTATGCTTGTCTTGTAGCTGCGCCAGTTCAAAGT 60
Db 60306 ATGAAAAAAGAAAGTATTGCTTTTACGCGCAGCTATTACATTAGTTGCTCCATTACAAAGT 60365
Qy 61 GTAGTATTTCACAAAACAAATAATAGTG---AAAGTCCTCGCACCCGATTTTAAAGATGGTCA 117
Db 60366 GTTGCATTGCTCATGAAATGATGGGGGACAGAGATTGGAGTTATTCGCGCGCTGCT 60425
Qy 118 GCTGAGGATAAGCATAAATAGAGGGGATTAATCTCTCAATTTGTGGATTGTAAATCGTCAATT 177
Db 60426 GCTGAAGATAAACATAAAGAGGCGTGAATCTCATTTATGGATTGTAAACCGTCAATT 60485
Qy 178 GACATCATGCTCGTAAATACACGATTCGTGATCCGATGAACCTGATTAATTAATAG 237
Db 60486 GATATTATGCTCGTAATACAAACACTTGTAAAAAAGATCGAGTTGCGACTATTAAATGAA 60545
Qy 238 TGGCGTGTGATTAGAAAATGGTATTATTCTCTGCTGATTACGAGAATCCCTTATTATGAT 297
Db 60546 TGGCGTACTGAGTAGAGAACGGTATTATTGCTGCTGCTATGATAAAATCCCTTATTATGAT 60605
Qy 298 AATAGTACATATGCTTCTCATTATGATCCGGATACTGGAAACAATATATTCCTTTT 357
```



```
Qy 418 AATCAAGATATGCACCAAGCAATCTTCTACTTAGGATTATCGCTTCATTATTAGGAGAT 477
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
601 AATAAGATATGAACAAGCAATCTTCTATTATTAGGATTATCTCTTCAATTATCTAGGGAT 660
Qy 478 GTGATCAGCAATCATGCAGCAACCTTACCAATCTTCTTATCCAAATGGGTTCCAT 537
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
661 GTAAACCAACCGATGCATGCAGCAACCTTACCAACCTTTCGTATCCCAAGGGTTCCAT 720
Qy 538 TCTAATACGAAATTTTGTGATACATAAATAAATAAATAAATAAATAAATAAATAAATAA 597
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
721 TCTAATATGAATCTTGTAGATACGATGAAGATTAATATAAGTAACGATCGAAT 780
Qy 598 GGATATTCGAATTTGAAGAGCAACCCAGAGATTTGATTAAGAGGACGCGGTAGCA 657
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
781 GGATATTCGAATTTGAAGAGGATTAAGATTAAGAGGATTTGATTAAGAGGACGCGGTAG 840
Qy 658 GCTAACAAGATTTCTCGGCTTGTGAACGATACGACAAAGATTTGTTGTTAAAGCA 717
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
841 GCGAACAAGATTTACGCTGGCAATTTGAATGATAATACGAAAGATTTGTTGTTAGAGCT 900
Qy 718 GCCGTATCTCAAGATATGCAGATAAATGGCGTGGCAAGTAACACCGGTGACAGGAAG 777
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
901 GCTGATACAGAATATGCAGATAAATGGCGCTGAGTTACACCAATGACAGGTAG 960
Qy 778 CGTTTAATGGAAGCCGCGGTTACAGCTGGTTATATTCATTTGTTGTTGATACGTAT 837
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
961 CGTTTAATGGAAGCCGCGGTTACAGCTGGTTATATTCATTTGTTGTTGATACGTAT 1020
Qy 838 GTAAATCGCTAA 849
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1021 GGAGATCGTTAA 1032
```

```
RESULT 9
BCPLC BCPLC 1622 bp DNA linear BCT 26-JUN-1997
LOCUS B. cereus gene for sphingomyelinase (EC 3.1.4.3).
```

```
DEFINITION X12711.1 GI:39439
ACCESSION X12711.1
VERSION 1
KEYWORDS phospholipase; phospholipase C; serine carboxypeptidase; sphingomyelinase.
```

```
SOURCE Bacillus cereus
ORGANISM Bacillus cereus
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
```

```
REFERENCE 1 (bases 1 to 1622)
AUTHORS Yamada,A., Tsukagoshi,N., Uchida,S., Sasaki,T., Makino,S., Nakamura,S., Little,C., Tomita,M. and Ikezawa,H.
TITLE Nucleotide sequence and expression in Escherichia coli of the gene coding for sphingomyelinase of Bacillus cereus
JOURNAL Eur. J. Biochem. 175 (2), 213-220 (1988)
MEDLINE 88296483
PUBMED 2841128
```

```
COMMENT Data kindly reviewed (25/4/89) by Tsukagoshi N.
FEATURES
Location/Qualifiers
1..1622
```

```
source
/organism="Bacillus cereus"
/mol_type="genomic DNA"
/strain="IAM1208"
/db_xref="taxon:1396"
/clone_lib="pUC18."
<1..358
/codon_start=2
/trans_table=11
```

```
CDs
/product="phosphatidylcholine-hydrolyzing phospholipase C"
/protein_id="CAA31213.1"
/db_xref="GI:757735"
/db_xref="GOA:P09598"
/db_xref="SWISS-PROT:P09598"
/translation="HAANFTNLSYPOGFHSKYENFVDIKNYKVTDGNGYWNWKGNTN PEDWTHGAIVAKQDYAGIVNDNTKDFVRAAVSQEYADKWRAEVPTMTGKRLMDAQR VTAGYIQLWFPTYGNR"
```

```
promoter 363..368
/note="35 region"
repeat_unit 366..381
/note="inverted repeat A"
promoter 386..392
/note="10 region"
repeat_unit 388..404
/note="inverted repeat A'"
RBS 422..426
/note="pot. ribosome binding site"
CDS 435..1436
/codon_start=1
/trans_table=11
/product="presphingomyelinase (AA -27 to 306)"
/protein_id="CAA31214.1"
/db_xref="GI:580758"
/db_xref="GOA:P11889"
/db_xref="SWISS-PROT:P11889"
/translation="WKGLKGLVSLGVLGALYSQAEASTNQNDTLKMTNHN YMLSTNLYPNWGQTERADLIGAADYIKQDVVILNEVPDNASDRLLGNLKKKEYNQI AVLRSGSEWEDKTLIGNYSSSTPEDGGVAVISRWPIAEIKIYVFAKGGCPDLSNKG FVYTKIKNDREHVIQTHLQABDSMCGTSPASVTRNQLKEIQDFIKNNIPNNEYVL IGGDMVNNKINAEENNDSYASMEFKTLNASVPSYTGHTATWDATNLSAKNFPDSPA EYLDYIIASKDHANPSYIENKVLQPKSPQWTVTSFQKYTYNDISDDYVPEATISMK"
435..515
/note="signal peptide (AA -27 to -1)"
516..1433
/product="sphingomyelinase (AA 1 - 306)"
1441..1453
/note="inverted repeat B"
repeat_unit 1459..1471
/note="inverted repeat B'"
ORIGIN
```

```
Query Match 30.1%; Score 255.6; DB 1; Length 1622;
Best Local Similarity 82.1%; Pred. No. 1e-42; Mismatches 0; Gaps 0;
Matches 294; Conservative 0; Indels 64;
Qy 492 GCATGCAGCAAACTTTACGAATCTTCTTCAATCCATGGGTTTCCATTTCTAAATACGAAA 551
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 GCATGGCGCAAACTTTACGAACCTTTCGTATCCCAAGGATTTCCATTTCTAAATACGAAA 60
Qy 552 TTTTGTGATACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 611
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 CTTTGTAGATACGATAAAGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 120
Qy 612 GAAAGGCAAAACCCAGAGATTGGATTGAAGGACGCGGTAGCAGCTTAACACAGATTA 671
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 GAAAGGTACGAATCCAGAGATTGGATTGAAGGACGCGGTAGTTCGGAACAGATTA 180
Qy 672 TCCTGGCGCTTGTGAACGATACGACAAAAGATTGGTTTGTAAAGCAGCGCGTATCTCAAGA 731
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 CGCTGGCATTGTAATGATATACGAAGATTGGTTTCGTAAGACGAGCGCTGATACAGA 240
Qy 732 ATATGCAGATAAATGGCGTGGGAAGTAACACCGGTGACAGAAAGCGTTTAATCGAAGC 791
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 ATATGCAGATAAATGGCGCGCTGGAAGTTACACCAATGACAGTAAGCGTTTAATCGATGC 300
Qy 792 GCAGCGCTTACAGCTGGTTATATTTCTTGTGGTTTGCATAGCTATGTAATCGCTAA 849
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 ACAACGTGTTACTGCTGGATACATTCAGCTTTGGTTTGTATACGTACGGAATTCGTTAA 358
```

```
RESULT 10
BACHEMOL BACHEMOL 280 bp DNA linear BCT 26-APR-1993
LOCUS B.cereus hemolysin gene, partial cds.
DEFINITION M35411
ACCESSION M35411
VERSION M35411.1 GI:143048
KEYWORDS hemolysin.
SOURCE Bacillus cereus
ORGANISM Bacillus cereus
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. .280
/organism="Bacillus cereus"
/mol_type="genomic DNA"
/db_xref="taxon:1396"
<1. .>280
/note="hemolysin"
/codon_start=2
/transl_table=11
/protein_id="AAA22524.1"
/db_xref="GI:143049"
/translacion="NSHMTVNRADIMSRNTTLVKQDRVALINWRTLENGIYAAD
YENPYDNSTFASHFYDPDNGKTYIPYAKQAKETGAKYFKLAGESYKXK"

Query Match 22.38; Score 189.4; DB 1; Length 280;
Best Local Similarity 79.94; Pred. No. 5.6e-29;
Matches 223; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 145 AACTCTCATTTGGATTGTAATCGTGCAATTGACATCATGTCGTGTAATACACGATT 204
Db AACTCTCATTTGGATTGTAATCGTGCAATTGACATCATGTCGTGTAATACACGATT 204

Qy 205 GTGATCCGAATGAACATGATTAATTAATGAGTGGCGTCTGATTTAGAAAATGGTATT 264
Db GTGATCCGAATGAACATGATTAATTAATGAGTGGCGTCTGATTTAGAAAATGGTATT 264

Qy 265 TATTCTGCTGATACAGAACTCTTATATGATTAATGATGATGATGATGATGATGATGAT 324
Db TATTCTGCTGATACAGAACTCTTATATGATTAATGATGATGATGATGATGATGATGAT 324

Qy 325 GATCCGGATCTGGAAACACATATATTCCTTTTGGAAACATGCAAAAGAAACAGCGCA 384
Db GATCCGGATCTGGAAACACATATATTCCTTTTGGAAACATGCAAAAGAAACAGCGCA 384

Qy 385 AATATTTTAACTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 423
Db AATATTTTAACTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 423

RESULT 11
AY150839
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
L.isteria seeligeri
L.isteria seeligeri
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE
AUTHORS
Schmid,M., Walcher,M., Bubert,A., Wagner,M., Wagner,M. and
Schleifer,K.H.
TITLE
Nucleic acid-based, cultivation-independent detection of Listeria
spp. and genotypes of L. monocytogenes
JOURNAL
MEDLINE
FEMS Immunol. Med. Microbiol. 35 (3), 215-225 (2003)
PUBMED
22535694
12648840
REFERENCE
2 (bases 1 to 825)
AUTHORS
Walcher,M., Schmid,M. and Wagner,M.
TITLE
Direct Submision
JOURNAL
Submitted (11-SEP-2002) Lehrstuhl fuer Mikrobiologie, Am Hochanger
4, Freising 85354, Germany
FEATURES
Location/Qualifiers
source
1. .825
```

```
/organism="Listeria seeligeri"
/mol_type="genomic DNA"
/db_xref="taxon:1640"
1. .>825
/gene="plcB"
1. .>825
/gene="plcB"
/note="lecithinase"
/codon_start=1
/transl_table=11
/product="phospholipase"
/protein_id="AAO19486.1"
/db_xref="GI:29373300"
/translacion="MKIKGILMLGLSALISVALSGADEASACGDSVKDOIAPHDION
KLPKLGWSAHSKNEINTHLWLNQAEKILAKHVTGAQLDLVRELKNTKEIATQGI
FLADKPNPYDKNTFLSHFNPYKTHKYIAGFPNAKDTGKYFNISIEEYDGNFEKA
FYNLGLAHVYTDISQPMHANNFTALSHFPGYHCAYENYVDYTFKQIFQNSABSEAKWF
CTDDVSEWFHENAKRAQADYPKIVNTIIKKSYITQGLSDSQKDRTWKKAARAATGKRL"

ORIGIN

Query Match 20.8%; Score 176.8; DB 1; Length 825;
Best Local Similarity 52.3%; Pred. No. 2e-26;
Matches 391; Conservative 0; Mismatches 357; Indels 0; Gaps 0;

Qy 36 TGCTTTAGCTCGCGCAGTTTCAAAGTGTAGTGTATTTGCAAAACAAATAATAGTGAAGTCC 95
Db TGCTTTAGCTCGCGCAGTTTCAAAGTGTAGTGTATTTGCAAAACAAATAATAGTGAAGTCC 95

Qy 96 TGCACCAATTTAAGATGGTCACTGAGGATAAGCATATGAGGGGATTAACCTCTCATTT 137
Db TGCACCAATTTAAGATGGTCACTGAGGATAAGCATATGAGGGGATTAACCTCTCATTT 137

Qy 138 ACCTAGTAAATAGGCTGGTCCAGCAGAGCATCTTCAAAAATGAAATTAATACATTT 155
Db ACCTAGTAAATAGGCTGGTCCAGCAGAGCATCTTCAAAAATGAAATTAATACATTT 155

Qy 156 GTGATTTGTAATCGTGCATATGACATCATGTCGTGTAATGACCAATGTTGTAATCCGAA 215
Db GTGATTTGTAATCGTGCATATGACATCATGTCGTGTAATGACCAATGTTGTAATCCGAA 215

Qy 198 GTGGTTGTTTAAATCAAGCAGAAAAAATCTTGCACCAATGTAACCTGGTCCCACTTGA 257
Db GTGGTTGTTTAAATCAAGCAGAAAAAATCTTGCACCAATGTAACCTGGTCCCACTTGA 257

Qy 216 TGAACCTGCAATTTAAATGAGTGGCGTCTGATTTAGAAAATGGTATTATTTCTGCTGA 275
Db TGAACCTGCAATTTAAATGAGTGGCGTCTGATTTAGAAAATGGTATTATTTCTGCTGA 275

Qy 258 TTTAGTAAAGAGATTTAAATAATTAATAAGAAATTTGCTCAAGGAATATTTGATGAGA 317
Db TTTAGTAAAGAGATTTAAATAATTAATAAGAAATTTGCTCAAGGAATATTTGATGAGA 317

Qy 276 TTAGAGAAATCTTATTTATGATTAATAGTATGCTTCTCACTTTTATGATCCGGATAC 335
Db TTAGAGAAATCTTATTTATGATTAATAGTATGCTTCTCACTTTTATGATCCGGATAC 335

Qy 318 TCATAAAACCCATATATGATGAAGAACTTTTATCGCATTTTATTAATCCAAAC 377
Db TCATAAAACCCATATATGATGAAGAACTTTTATCGCATTTTATTAATCCAAAC 377

Qy 336 TGGACCAACATATATTCCTTTTGGGAAACATGCAAAAGAAACAGCGGCAAAATATTTAA 395
Db TGGACCAACATATATTCCTTTTGGGAAACATGCAAAAGAAACAGCGGCAAAATATTTAA 395

Qy 378 TCATAAAACCTCATAGCGGATTTCTTAATGCTAAAGATACAGGCACCTAAATACCTTAA 437
Db TCATAAAACCTCATAGCGGATTTCTTAATGCTAAAGATACAGGCACCTAAATACCTTAA 437

Qy 396 CTTGCTGCTCAAGCATACCAAAATCAAGATATGACGCAAGCATTTCTTCTAGTTAGGATT 455
Db CTTGCTGCTCAAGCATACCAAAATCAAGATATGACGCAAGCATTTCTTCTAGTTAGGATT 455

Qy 438 TATTTCCATTGAAGAGTATCAAGATGGAACCTTTGAAAAGCATTTTAACTTAGGATT 497
Db TATTTCCATTGAAGAGTATCAAGATGGAACCTTTGAAAAGCATTTTAACTTAGGATT 497

Qy 456 ATCGCTTCATTTATTTAGGAGATGTAATCAGCAATGATGATGATGATGATGATGATGAT 515
Db ATCGCTTCATTTATTTAGGAGATGTAATCAGCAATGATGATGATGATGATGATGATGAT 515

Qy 498 AGCAATACATTTATATACGGATATTAGTCAGCGCATGATGCAATAAATTAACCTTTACG 557
Db AGCAATACATTTATATACGGATATTAGTCAGCGCATGATGCAATAAATTAACCTTTACG 557

Qy 516 TTCTTATCCAATGGGTTTCCATTTCTAAATACCAAAATTTGTTGATACAAATAAATAA 575
Db TTCTTATCCAATGGGTTTCCATTTCTAAATACCAAAATTTGTTGATACAAATAAATAA 575

Qy 558 TTCTCACCCTGATGATATCACTGCGCATATGAAAATTTATGGAACAATTTAGGCAAT 617
Db TTCTCACCCTGATGATATCACTGCGCATATGAAAATTTATGGAACAATTTAGGCAAT 617

Qy 576 CTATATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 635
Db CTATATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 635

Qy 618 TTTCCAGCGCTCCGAGAGTTCAGAAAGCAAGTGGTCTCTGATGATGATGATGATGATG 677
Db TTTCCAGCGCTCCGAGAGTTCAGAAAGCAAGTGGTCTCTGATGATGATGATGATGATG 677

Qy 636 GATTGAAGGAGCAGCGGTAGCAGCTTAAACAGATTTCTGCGCTTTGTAACGATACGAC 695
Db GATTGAAGGAGCAGCGGTAGCAGCTTAAACAGATTTCTGCGCTTTGTAACGATACGAC 695

Qy 678 GTTCCATGAAATGCAAAAGAGCAGCGGATTTACCCAAAATAGTATTAATCTATAAT 737
Db GTTCCATGAAATGCAAAAGAGCAGCGGATTTACCCAAAATAGTATTAATCTATAAT 737

Qy 696 AAAAGATGGTTGTTGAAAAGCAGCGGATTTCTCAAGAAATATGCAAGAAATAGGCGT 755
Db AAAAGATGGTTGTTGAAAAGCAGCGGATTTCTCAAGAAATATGCAAGAAATAGGCGT 755

Qy 738 TAAAAAATCTTATATACAGGGCTTTTCAAGTCTTCAAAAAGATAGAACTTTGGAAGAAC 797
Db TAAAAAATCTTATATACAGGGCTTTTCAAGTCTTCAAAAAGATAGAACTTTGGAAGAAC 797

Qy 756 AGTAACACCGGTGACAGGAAAGCGTTTA 783
```

```
Db 798 AGTAAGCGCTGCTACTCGAAGAGACTA 825
|||||
RESULT 12
AX413749
LOCUS AX413749 870 bp DNA linear PAT 02-SEP-2002
DEFINITION Sequence 740 from Patent WO0228891.
ACCESSION AX413749
VERSION AX413749.1 GI:21446206
KEYWORDS
SOURCE Listeria monocytogenes EGD-e
ORGANISM Listeria monocytogenes EGD-e
REFERENCE 1
AUTHORS Kunst, F. and Glaser, P.
TITLE Listeria innocua, genome and applications
JOURNAL Patent: WO 0228891-A 740 11-APR-2002;
INSTITUT PASTEUR (FR); CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (FR)
FEATURES
source Location/Qualifiers
1..870
/organism="Listeria monocytogenes EGD-e"
/mol_type="unassigned DNA"
/db_xref="taxon:169963"
ORIGIN
Query Match 18.2%; Score 154.4; DB 6; Length 870;
Best Local Similarity 53.1%; Pred. No. 8.2e-22;
Matches 329; Conservative 0; Mismatches 291; Indels 0; Gaps 0;
Qy 98 CACCGATTTTAAAGATGGTCAGCTGAGGATAAGCAATAATGAGGGGATTAACTCTCATTTGT 157
|||
Db 140 CACATAAACTTAGTTGGTCCGGGATAACCGACAAATACTGACGTAATAACGCACTATT 199
|||
Qy 158 GGATTGTAATCGTGCAATTGACATCATGCTCGTAATACACGATTTGTAATCGAATG 217
|||
Db 200 GGCTTTTAAACAAGCGGAAAAATACTAGCTAAAGATGTAAATCATATGCGAGCTAAT 259
|||
Qy 218 AAATGCAATTAATAAGTGGCGTGTGATTTAGAAAAATGGTATTTTCTGCTGATT 277
|||
Db 260 TAATGAATGAATTAATAAATTCGATAAACAATAGCTCAAGGAATATATGATCGGATC 319
|||
Qy 278 ACGAAGATCCTTATATGATAATAGTACATATGCTTCTCACTTTTATGATCGGATACG 337
|||
Db 320 ATAAAAATCCATATATGATAGTACATTTTATCTCAATTTTATATATCTCTGATAGAG 379
|||
Qy 338 GAACAACATATATTCCTTTTGGAAACATGCAAAAGAAACAGCGCCAAATATTTTAACC 397
|||
Db 380 ATAATACCTTATTTGCGGGTTTGTCTAATGCGAAAAATTAACAGGACAAAGTATTTCAATC 439
|||
Qy 398 TTGCTGGTCAAGCATACCAAAATCAAGATATGCGACAGCAATTCCTTCTACTTAGGATTAT 457
|||
Db 440 AATCGGTGACTGATTTACGAGAGGAAATTTGACACAGCGTTTATATAATTAGCCCTAG 499
|||
Qy 458 CGCTTCATTTATAGGAGATGTGAATCAGCAATGCAATGCGACGAACTTTTACGAATCTTT 517
|||
Db 500 CAATCCATTTATATACGGATATTAGTCAACCTATGCGACGCCAAATAATTTTACCGCAATAT 559
|||
Qy 518 CTTATCCAATGGGTTTCCATTTCTAAATACGAAAAATTTTGTGATACATAAAAAATAACT 577
|||
Db 560 CATACCTTCCAGGCTACCACTGTGCATATGAAAATTTACGTAGATACCAATTAACACAAT 619
|||
Qy 578 ATATTGTTTTCAGATAGCAATGGATATTGGAATTTGGAAGGACCAACCCAGAAAGTTGGA 637
|||
Db 620 ATCAAGCAACGGAAGACATGTTAGCAAAAAGATTTTGTCTCAGATGACGTGAAGACTGCG 679
|||
Qy 638 TTGAAGGAGCAGCGGTAGCAGCTAAACAAGATTTATCTGGCGTTTGTGAACGATACGACAA 697
|||
Db 680 TCTATGAAAATGCGAAAAGGGCGAAGCGGACTACCCGAAAATAGTCAATGCGAAAACCTA 739
|||
Qy 698 AAGATTGGTTTGTAAAAGCA 717
|||
```

```
Db 740 AAAAATCATATTTTAGTAGGA 759
|||||
RESULT 13
AX415655
LOCUS AX415655 870 bp DNA linear PAT 02-SEP-2002
DEFINITION Sequence 2646 from Patent WO0228891.
ACCESSION AX415655
VERSION AX415655.1 GI:21448112
KEYWORDS
SOURCE Listeria monocytogenes EGD-e
ORGANISM Listeria monocytogenes EGD-e
REFERENCE 1
AUTHORS Kunst, F. and Glaser, P.
TITLE Listeria innocua, genome and applications
JOURNAL Patent: WO 0228891-A 2646 11-APR-2002;
INSTITUT PASTEUR (FR); CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (FR)
FEATURES
source Location/Qualifiers
1..870
/organism="Listeria monocytogenes EGD-e"
/mol_type="unassigned DNA"
/db_xref="taxon:169963"
ORIGIN
Query Match 18.2%; Score 154.4; DB 6; Length 870;
Best Local Similarity 53.1%; Pred. No. 8.2e-22;
Matches 329; Conservative 0; Mismatches 291; Indels 0; Gaps 0;
Qy 98 CACCGATTTTAAAGATGGTCAGCTGAGGATAAGCAATAATGAGGGGATTAACTCTCATTTGT 157
|||
Db 140 CACATAAACTTAGTTGGTCCGGGATAACCGACAAATACTGACGTAATAACGCACTATT 199
|||
Qy 158 GGATTGTAATCGTGCAATTGACATCATGCTCGTAATACACGATTTGTAATCGAATG 217
|||
Db 200 GGCTTTTAAACAAGCGGAAAAATACTAGCTAAAGATGTAAATCATATGCGAGCTAAT 259
|||
Qy 218 AAATGCAATTAATAAGTGGCGTGTGATTTAGAAAAATGGTATTTTCTGCTGATT 277
|||
Db 260 TAATGAATGAATTAATAAATTCGATAAACAATAGCTCAAGGAATATATGATCGGATC 319
|||
Qy 278 ACGAAGATCCTTATATGATAATAGTACATATGCTTCTCACTTTTATGATCGGATACG 337
|||
Db 320 ATAAAAATCCATATATGATAGTACATTTTATCTCAATTTTATATATCTCTGATAGAG 379
|||
Qy 338 GAACAACATATATTCCTTTTGGAAACATGCAAAAGAAACAGCGCCAAATATTTTAACC 397
|||
Db 380 ATAATACCTTATTTGCGGGTTTGTCTAATGCGAAAAATTAACAGGACAAAGTATTTCAATC 439
|||
Qy 398 TTGCTGGTCAAGCATACCAAAATCAAGATATGCGACAGCAATTCCTTCTACTTAGGATTAT 457
|||
Db 440 AATCGGTGACTGATTTACGAGAGGAAATTTGACACAGCGTTTATATAATTAGCCCTAG 499
|||
Qy 458 CGCTTCATTTATAGGAGATGTGAATCAGCAATGCAATGCGACGAACTTTTACGAATCTTT 517
|||
Db 500 CAATCCATTTATATACGGATATTAGTCAACCTATGCGACGCCAAATAATTTTACCGCAATAT 559
|||
Qy 518 CTTATCCAATGGGTTTCCATTTCTAAATACGAAAAATTTTGTGATACATAAAAAATAACT 577
|||
Db 560 CATACCTTCCAGGCTACCACTGTGCATATGAAAATTTACGTAGATACCAATTAACACAAT 619
|||
Qy 578 ATATTGTTTTCAGATAGCAATGGATATTGGAATTTGGAAGGACCAACCCAGAAAGTTGGA 637
|||
Db 620 ATCAAGCAACGGAAGACATGTTAGCAAAAAGATTTTGTCTCAGATGACGTGAAGACTGCG 679
|||
Qy 638 TTGAAGGAGCAGCGGTAGCAGCTAAACAAGATTTATCTGGCGTTTGTGAACGATACGACAA 697
|||
Db 680 TCTATGAAAATGCGAAAAGGGCGAAGCGGACTACCCGAAAATAGTCAATGCGAAAACCTA 739
|||
Qy 698 AAGATTGGTTTGTAAAAGCA 717
|||
Db 740 AAAAATCATATTTTAGTAGGA 759
|||||
```

```
RESULT 14
LOCUS LISACTLHDH
DEFINITION Listeria monocytogenes lecithinase, surface antigen. BCT 26-APR-1993
cgs, complete cds, (plcB) gene complete cds, (ldh) gene complete
cgs.
ACCESSION M82881
VERSION 1
KEYWORDS GI:149644
SOURCE lactate dehydrogenase; lecithinase; surface antigen.
ORGANISM Listeria monocytogenes
REFERENCE 1 (bases 1 to 5648)
AUTHORS Vasquez-Boland J.A., Kocks, C., Dransil, S., Ohayon, H., Geoffroy, C.,
Mengaud, J. and Cossart, P.
TITLE Nucleotide sequence of the lecithinase operon of Listeria
monocytogenes and possible role of lecithinase in cell-to-cell
spread
JOURNAL Infect. Immun. 60 (1), 219-230 (1992)
MEDLINE 92104678
PUBMED 1309513
COMMENT Original source text: Listeria monocytogenes (strain L028) DNA.
FEATURES
source Location/Qualifiers
1..5648
/organism="Listeria monocytogenes"
/mol_type="genomic DNA"
/strain="L028"
/db_xref="taxon:1639"
228..2147
/gene="actA"
228..2147
/gene="actA"
/codon_start=1
/transl_table=1
/protein_id="AAA25270.1"
/db_xref="GI:149645"
/translation="MGLNRFMRMMVVFITANCITINPDIIFAATSDSSSLNTDWE
EKTREQSVNTGPRYETAREVSRDIELEKSNKVRNTKADLIAMLEKAKGPN
INNNSQETENAAINEEASGADPAIQVERHPGLPSDAAIEKKRAIASSELE
SITYDPKPTKVKKVKAKESVADASESLDSMQSADSSPOPLKAKQPPFPVKFK
INDAGWVRDKIDENPEVKAIVDKSAGLIDQLTKKSEVNASDFPPTDEBELR
ALPETPMLGNAPATSPSPSEFPPTDEBELRALPETPMLGNAPATSPSPSE
FPPTDELEIIRITASLDSLSTFSGDLSLRNAINRHSQNSDFPPIPTSEELNGR
GSRPSEFSSNGDFTDENSETTEEIIDLADLRDGTGKHSNAGLPLNPPAS
SPVSLSPKVSISAPALISDITKTPFPNSQPLNPNKTTTWTVKTPVKTAP
KLALPATKQPTVLRNKTPPIEKOAKTNKSNMPSLPVIOKEATESKEEMKPT
EKMVESESANNANGKRSAGIEBGLKIAKSAEDERAKKEPFGNHTLLILAMLAIGVF
SLGAFIKIILRNKN"
2183..3052
/gene="plcB"
2183..3052
/gene="plcB"
/codon_start=1
/transl_table=1
/protein_id="AAA25270.1"
/db_xref="GI:149646"
/translation="MKFAKVVGLGMLIASVLVFPVTIKANACDEYLTQPAAPHIDS
KLPHKLSANPNTDNTVHYLWFLQAEKILAKDVNHNMEANLMLKPKDQIAQI
YDADHKNPYDTSTFLSHFYNDPNTYLPGFANAKITGAKYFNQSVTDYRKGFTDA
FYKGLAIHYTIDISQPMHANNFTAISYPQYHCAENYVDTIKNYQATEDMVAKRF
CSDVDKWLVENAKRAKADVPKIVNANTKSYLVGNSEWKDITVEPTGARLRDSQTL
AGFLFELFSSKKTNE"
3124..3540
/misc_feature /note="ORF/x/y"
3582..4043
/gene="ORF2"
/codon_start=1
/transl_table=1
/protein_id="AAA25271.1"
/db_xref="GI:149647"
/translation="MKMLKIGTAVLFVIMINAVMLVACGDKEETKTYTLTSQNGVDSKLT
YTKGDKVTKOTATMSYASIGVASKEDAEMKMKATSDKFGQIDGLKELEYKDDKA
4050..4093
complement(4094..4426)
/note="ORFB"
/codon_start=1
/transl_table=1
/protein_id="AAA25272.1"
/db_xref="GI:149648"
/translation="MIVTSPNIGKQIIEYKKIVFGEVITGVNFMKDIGAGLRNFFG
GRSGYRDELINAREEALIREMQRAKIGANAVIGVDIDYEVLGADNMLMVTASGTA
VVIEAQDY"
complement(4493..5167)
/note="ORPA"
/codon_start=1
/transl_table=1
/protein_id="AAA25273.1"
/db_xref="GI:149649"
/translation="MITFDQLDTLQALENPNTIKIPRNHGPCDSDLYGLKIGDLKK
IIRREKLTQNHSLAVKLIENNSDLIYGLLAINPNKITTQIEKKNIAFRETWISOLT
FGLASIVSKRDDALLFAKTWIESDYDLTKSMHQIYSEHINNLEPEATLQRAKETIQ
TETNTRYVMNGFIITCGIYKDDLHEKAMEAKSVGKVNHLNLTACKVDPDAISYIEK
ARNRTK"
complement(5244..5648)
/gene="ldh"
/codon_start=1
/transl_table=1
/protein_id="AAA25274.1"
/db_xref="GI:149650"
/translation="EPPANSHHTVGGIPTEWISBDEQGMDFIVSVRDAAYEIIINK
KGATFYGVAALARIYKAILNENNAIPLSVLDLGHYGMNDIYIGAPVAVNRQGVRI
VENLNDKEKEQMKNSADTLKVLDDAMKQID"
complement(5247..5648)
/gene="ldh"
ORIGIN
Query Match 18.28; Score 154.4; DB 1; Length 5648;
Best Local Similarity 53.1%; Pred. No. 6.8e-22;
Matches 329; Conservative 0; Mismatches 291; Indels 0; Gaps 0;
Qy 98 CACCGATTTTAAAGATGTCAGCTGAGGATAAGCATATGAGGGGATTAACCTCTCATTTCT 157
Db 2322 CACATAAACTTAGTTGTCGCGGATAACCCGACAAATAGTACGTAATAAGCCTATT 2381
Qy 158 GGATTTGAATCTGCAATGACATCTGCTGTAATCAACGATGGAATCCGAATG 217
Db 2382 GGCTTTTAAACAGCGGAAAAATACTAGCTAAAGATGATAATCATATCGAGCTAAT 2441
Qy 218 AAATCGCATTTTAAATGAGTGGCGTGTGATTTAGAAAATCGTATTTATCTCTGCTGATT 277
Db 2442 TAATGAATGAATTTAAAAAATTCGATAACAATAGCTCAAGAAATATATGATCGGATC 2501
Qy 278 ACAGAAATCCTTATATGATAATAGTACATATGCTTCTCAGCTTTTATGATCCGATACCTG 337
Db 2502 ATAAAAATCCATATATGATAGTACTAGTACATTTTATCTCTCATTTTATATCTGATAG 2561
Qy 338 GAACACATATATTCCTTTTGGAAACATGCAAGAAAGAACAGCCGCAAAATATTTTAACC 397
Db 2562 ATAAATCTTATTTGCGGGTTTGTGTAATGCAAAATTAACAGAGCAAAAGTATTTCAATC 2621
Qy 398 TTGCTGGTCAAGCATACCAAAATCAAGATATGACAGAGCATTTCTTCTACTTAGGATTAT 457
Db 2622 AATCGGTGACTGATTACCGAGAAAGGAAATTTGACACAGCGTTTTTATAAATAGGCCTAG 2681
Qy 458 CGCTTCATTATTAGGAGATGTAATCAGCAATGATGATGATGATGATGATGATGATGATGAT 517
Db 2682 CAATCCATTATTATAGGATATTAGTCAACCTATGACCGCAATAATTTTACCGCAATAT 2741
Qy 518 CTTATCCAATGGTTTCCATTCTAAATACGAAAAATTTTGTGTGATACAAATAAAAAATAACT 577
Db 2742 CATACCTCCAGGCTACCACTGTGCATATGAAAAATACGTAGATACCATTAACACATATT 2801
Qy 578 ATATTGTTTCAGATAGCATGATATGGATTTGGAAGGAGGAAACCCAGAGATTGGA 637
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 02:14:50 ; Search time 483 Seconds
(without alignments)
8618.563 Million cell updates/sec

Title: PCT-US03-12556-1

Perfect score: 849
Sequence: 1 atgaaaagaagattagc.....atagctatgtaaatcgctaa 849

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	849	100.0	849	16	US-10-421-654-1
2	559.2	65.9	1192	16	US-10-369-493-40434
3	529	62.3	852	16	US-10-421-654-3
4	505.8	59.6	843	16	US-10-421-654-5
5	154.4	18.2	870	16	US-10-398-221-740
6	154.4	18.2	870	16	US-10-398-221-2646
7	87.4	10.3	1422	16	US-10-421-654-31
8	85.8	10.1	1422	16	US-10-421-654-97
9	76.8	9.0	1422	16	US-10-421-654-89
10	72	8.5	1422	16	US-10-421-654-81
11	57.4	6.8	4562	16	US-10-398-221-3772
12	43.4	5.1	9821	8	US-08-781-986A-470
13	43.4	5.1	9821	13	US-10-329-624-470
14	43.2	5.1	837	10	US-09-966-521-21

c 15	43.2	5.1	837	15	US-10-429-094-21	Sequence 21, Appl
c 16	43.2	5.1	5924	8	US-08-781-986A-130	Sequence 130, App
c 17	43.2	5.1	5924	13	US-10-329-624-130	Sequence 130, App
c 18	43	5.1	8170	15	US-10-240-453-132	Sequence 132, App
c 19	41.8	4.9	2000	15	US-10-260-238-1941	Sequence 1941, App
c 20	41.8	4.9	6227	15	US-10-311-455-1559	Sequence 1559, App
c 21	41.8	4.9	6227	17	US-10-240-589C-69	Sequence 69, Appli
c 22	41.6	4.9	2319	14	US-10-067-385-7	Sequence 7, Appli
c 23	41.6	4.9	2911	13	US-10-158-844-199	Sequence 199, App
c 24	41.6	4.9	6360	10	US-09-769-744A-27	Sequence 27, Appl
c 25	41.6	4.9	6423	13	US-10-282-122A-37486	Sequence 37486, A
c 26	41.6	4.9	495269	16	US-10-398-221-8	Sequence 8, Appli
c 27	41.6	4.9	301208	16	US-10-398-221-2058	Sequence 2058, Ap
c 28	41.2	4.9	3673778	15	US-10-312-841-2	Sequence 2, Appli
c 29	41	4.8	4985	15	US-10-094-240-10	Sequence 10, Appl
c 30	41	4.8	4985	15	US-10-056-405-10	Sequence 10, Appl
c 31	40.4	4.8	541	13	US-10-424-599-52877	Sequence 52877, A
c 32	40.2	4.7	1014	13	US-10-282-122A-31009	Sequence 31009, A
c 33	40	4.7	3381	9	US-09-815-242-4318	Sequence 4318, Ap
c 34	40	4.7	3477	9	US-09-815-242-8424	Sequence 8424, Ap
c 35	39.8	4.7	1353	15	US-10-349-680-131	Sequence 131, App
c 36	39.8	4.7	2130	13	US-10-282-122A-40566	Sequence 40566, A
c 37	39.6	4.7	3673778	15	US-10-312-841-1	Sequence 1, Appli
c 38	39.4	4.6	5338	13	US-10-221-613-241	Sequence 241, Appl
c 39	39.2	4.6	13038	15	US-10-311-455-1247	Sequence 1247, Ap
c 40	39.2	4.6	60461	16	US-10-341-434-82	Sequence 82, Appl
c 41	39.2	4.6	150351	17	US-10-322-281-453	Sequence 453, App
c 42	39	4.6	835	15	US-10-029-386-22786	Sequence 22786, A
c 43	39	4.6	16217	15	US-10-311-455-597	Sequence 597, App
c 44	39	4.6	640681	9	US-09-790-988-1	Sequence 1, Appli
c 45	38.8	4.6	2583	17	US-10-441-926-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-10-421-654-1
; Sequence 1, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
US-10-421-654-1

Query Match 100.0%; Score 849; DB 16; Length 849;
Best Local Similarity 100.0%; Pred. No. 5.6e-196;
Matches 849; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAAGAAAGTATTAGCTAGCAGCTAGTGGTTTGTAGTGCAGCTTCAAGT 60
Db 1 ATGAAAAAGAAAGTATTAGCTAGCAGCTAGTGGTTTGTAGTGCAGCTTCAAGT 60
QY 61 GTAGTATTTCACAAACAAATAATAGTGAAGTCTGCACCGATTTTAAAGTGTGACCT 120
Db 61 GTAGTATTTCACAAACAAATAATAGTGAAGTCTGCACCGATTTTAAAGTGTGACCT 120


```
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; US-10-421-654-3

Query Match      62.3%; Score 529; DB 16; Length 852;
Best Local Similarity 76.4%; Pred. No. 3.2e-118; Mismatches 200; Indels 0; Gaps 0;
Matches 649; Conservative 0; Mismatches 200; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGAAAGTATTAGCACTAGCAGCTATGGTTGCTTTAGCTGCGCCAGTTCAAAAGT 60
Db 1 ATGAAAAAGAAATTTAGCTATAGCTCCGTAATGCTTTTAAACAGCTCCTATCCAAAGT 60

Qy 61 GTAGTATTGTCACAAACAAATTAATAGTGAAGTCTCTGACCGAATTTTAAAGATGGTCAGT 120
Db 61 GTGGCGTTTGGCGATGAAATGGTCAACCAAGATCCACCAATTTGCTCTAAAGTGGTCAGCA 120

Qy 121 GAGGATAGCAATAGCAGGGGATTAACCTCATTTTGTGATTTGAATCGTCAATTGAC 180
Db 121 GAATCTATACATAATGAAGGAGTAGTCTCATTTTATGATTTGTAAGGATTTGAACAGAGCCATTGAT 180

Qy 181 ATCATGCTCTCGTAATACAAACGATTTGTAATCCGAATGAACTGCAATTTAATTAATGAGTGG 240
Db 181 ATTATGTCCTCAAAATACGCTGTTGTAAGCAAAATGACAGAGCTCTATTAAATGAATGG 240

Qy 241 CGTGTGATTAGAAAATGGTATTATTCTGCTGATTAGAGAAATCCCTTATTATGATAAT 300
Db 241 CGTACGGATCTAGAGAAAGGCAATTTACTCTGGGATTTATGAAACCCCATCTATGATAAT 300

Qy 301 AGTACATATGCTTCTCATTATGATCCGGATCTGGAACCAACATATATTCCTTTTGGG 360
Db 301 TCCACATTTGCTTCTCATTATGATCCGGATCTGGAACCAACATATATTCCTTTTGGG 360

Qy 361 AAACATGCAAAAGAAACAGGGCCAAAATATTTTAACTTGTCTGCTCAAGCATACCAAAAT 420
Db 361 AAACAGCAAAAGCAACAGGAGCAAAATATTTTAAATGATGTTGTAAGCTTATCAAAAT 420

Qy 421 CAAGATATGCAAGCAAGATTTCTTACTTATGAGTATTCGTTTCAATTTATGAGAGATGG 480
Db 421 AAAGATCTGAAAACGCAATCTTTTATTATGAGTATTCATCTCACTATTTAGGGGATGTC 480

Qy 481 AATGAGCCAAATGCAATGCAAGCAATTTTACGAATCTTTCTATCCAAATGGGTTTCAATCT 540
Db 481 AACCAACCAATGCAATGCAAGCAATTTTACTAATATTTTGCATCTTATTTGGCTTCCACTCA 540

Qy 541 AAATACGAAAATTTTGTGATACAAATAAAAATACTATATTGTTTCAGATAGCAATGGA 600
Db 541 AAATATGAAAATTTTGTGATACAGTGAAGCAATATATGAGTAAACGATGGAATGGC 600

Qy 601 TATTGGAATTTGAAAGGCAACCCAGAGATTTGGATTGGAAGGAGCAGCGGTAGCAGT 660
Db 601 TATTGGAATTTGCAAGTGAATTTCCAGAGATTTGGATTGGAAGGAGCAGCGGTAGCAGT 660

Qy 661 AAACAGATTTATCTGGGCTTTGTGAACGATACGACAAAAGATTTGTTGTAAGAGCGCC 720
Db 661 AAACAGATTTATCTGGGCTTTGTGAACGATACGACAAAAGATTTGTTGTAAGAGCGCC 720

Db 661 AAAGCTGATTTTCCATCAATTTGTTAATGATAGACAAAATTTGGTTCTCTAAAGCAGCT 720
Qy 721 GTATCTCAAGATATGCAATAAATGGGTGGGAGTAACACCGGTGACAGGAAAGCT 780
Db 721 GTATCAAGAAGTCTGCTGATATAAATGGGTGAGAGTAACACCGATACAGGAAAGCT 780
Qy 781 TTAATGGAAGCGCAGCGGTACAGCTGGTTATATTTCAATTTGTTGTTGATACGTATGTA 840
Db 781 TTAATGGAAGCGCAGCGGTACAGCTGGTTATATTTCAATTTGTTGTTGATACGTATGTA 840

Qy 841 AATCGCTAA 849
Db 841 AATAACAAA 849

RESULT 4
US-10-421-654-5
; Sequence 5, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 843
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; US-10-421-654-5

Query Match      59.6%; Score 505.8; DB 16; Length 843;
Best Local Similarity 75.7%; Pred. No. 1.4e-112;
Matches 643; Conservative 0; Mismatches 197; Indels 9; Gaps 1;

Qy 1 ATGAAAAAGAAAGTATTAGCACTAGCAGCTATGGTTGCTTTAGCTGCGCCAGTTCAAAAGT 60
Db 1 ATGAAAAAGAAATTTAGCTATAGCTTCTGTAATTTGCTTTTAAACAGCTCCTATTCNAAGT 60

Qy 61 GTAGTATTGTCACAAACAAATAATAGTGAAGTCTCTGACCGAATTTTAAAGATGGTCAGT 120
Db 61 GTGGCGTTTGGCGATGAATCTGATGGGCTATTGC-----TTTAAAGATGGTCAGCG 111

Qy 121 GAGGATAGCAATAGCAGGGGATTAACCTCATTTTGTGATTTGAATCGTCAATTGAC 180
Db 112 GAATCTGATCAATAATGAAGGAGTAGTCTCATTTATGATTTGTAACAGAGCAATTTGAT 171

Qy 181 ATCATGCTCTCGTAATACAAACGATTTGTAATCCGAATGAACTGCAATTTAATTAATGAGTGG 240
Db 172 ATTATGTCCTCAAAATACGCTGTTGTAAGCAAAATGACAGAGCTCTATTAAATGAATGG 231

Qy 241 CGTGTGATTAGAAAATGGTATTATTCTGCTGATTACGAAATCTTATTTATGATAAT 300
Db 232 CGTACGAATTTGGAGGAAGGTATTATTCTGAGATTTATAAAACCCCATCTATGATAAT 291

Qy 301 AGTACATATGCTTCTCATTATGATCCGGATCTGGAACCAACATATATTCCTTTTGGG 360
Db 292 TCCACATTTGCTTCACTTCTATGATCTGATTCAGAAAACAGTATATTTCCATTTGCT 351

Qy 361 AAACATGCAAAAGAAACAGGGCCAAAATATTTTAACTTGTCTGCTCAAGCATACCAAAAT 420
Db 361 AAACAGCAAAAGCAACAGGAGCAAAATATTTTAAATGATGTTGTAAGAGCTTATCAAAAT 411
```

```
QY 421 CAAGATATGACGACGATCTTCTTCTACTTAGGATATCGCTTCATTAATTTAGGAGATGTG 480
D 412 AAAGATCTGAAAAATGCATCTTTTATTTAGGATATCACTTCATTAATTTAGGAGATGTG 471
QY 481 AATCGCCAAATGCATGCGAGCAAACTTTACGAACTCTTTCTTATCCATGGGTTTCCATCTCT 540
D 472 AATCAACCAATGCATGCGAGCAAACTTTTACTAACATTTTCGCATCCATTTGGGCTTCCACTCA 531
QY 541 AAATACGAAAAATTTGTTGATACAAATAAAAAATAACTATATTTGTTTCAGATAGCAATGGA 600
D 532 AAATATGAAACTCTGTTGATACAGTGAAGCAAAATATAGATTAACAGATGGAGATGGC 591
QY 601 TATTTGAAATGGAAGGAGCAAAACCCAGAGATTTGATTTGAAGGAGCAGCGGTAGCAGCT 660
D 592 TATTTGAAATGGAAGGAGCAAAACCCAGAGATTTGATTTGATGATCAGCATCAGCATCAGCAGCA 651
QY 661 AAACAGAAATATCTGGCGGTTGTGAAACGATACGACAAAGAAATTTGTTTGAAGGAGCAGCC 720
D 652 AAAGCTGATTTCCCATCAATTTGTTAAATGATTAATACGAAAGTTGGTTCTTAAAGCAGCG 711
QY 721 GTATCTCAAGAAATATGCAGATAAATGGCGTGGGAAGTAAACACCGGTGACAGGAAAGCGT 780
D 712 GTATCACAAGACTCTGCTGACAAATGGCGTGTCTGAAGTAAACACCGGTGACAGGAAAGCGT 771
QY 781 TTAATGGAAGGAGCAGCGGTTTACAGCTGGTTATATTTGATTTGTTGTTGATACGATGTA 840
D 772 TTAATGGAAGCAGCAGCGTATTACAGCTGGATATATTTCAATTTATGTTTGTGATACGATACG 831
QY 841 AATCGCTAA 849
D 832 AATAACAAA 840
```

RESULT 5

US-10-398-221-740

; Sequence 740, Application US/10398221

; Publication No. US20040018514A1

; GENERAL INFORMATION:

; APPLICANT: KUNST, Frederik

; APPLICANT: GLASER, Philippe

; TITLE OF INVENTION: Listeria innocua, genome and applications

; FILE REFERENCE: 344 702 - US

; CURRENT APPLICATION NUMBER: US/10/398,221

; PRIOR FILING DATE: 2003-03-27

; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061

; PRIOR FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: FR 00/12 697

; PRIOR FILING DATE: 2000-10-04

; NUMBER OF SEQ ID NOS: 4025

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 740

; LENGTH: 870

; TYPE: DNA

; ORGANISM: Listeria monocytogenes-EGD

US-10-398-221-740

Query Match 18.2%; Score 154.4; DB 16; Length 870;
Best Local Similarity 53.1%; Pred. No. 3.3e-27;
Matches 329; Conservative 0; Mismatches 291; Indels 0; Gaps 0;

```
QY 98 CACCGATTTTAAAGATGGTCAGCTGAGGATTAAGCATAAATGAGGGGATTAACCTCTCATTTGT 157
D 140 CACATAAACTTAGTTGGTCCGGGATAACCCGACAAATACTGACGTAAATACGCACTATT 199
QY 158 GGAATGTAATCGTGAATTTGACATCACTGCTCGTAATACACGATTTGTAATCCGAATG 217
D 200 GGCCTTTTAAACAGCGGAAAAATACTAGCTAAAGATGTAATCATATGCGAGCTAATT 259
QY 218 AAATGCAATTTAAATGAGTGGCGTCTGATTTAGAAAAATGGTATTTATTTCTGCTGATT 277
D 260 TAAATGAATGAACCTTAAAAAATTCGATAAAACAAATAGCTCAAGGAATATATGATCGGATC 319
QY 278 ACGAATCCCTTATTATGATATAATAGTACATATGCTTCTCACTTTTATGATCCGGATCTG 337
```

```
D 320 ATAAAAATCCATATTAATGATACATAGTACATTTTATCTCATTTTATAATCCTCATAGAG 379
QY 338 GAACCAACATATATCTCTTTTTCGGAACATCGCAAAAGAAACAGGCGCAAAATATTTTAACC 397
D 380 ATAATACTTATTTGCCGGGTTTGGCTAATGCGAAATAACAGGAGCAAGATTTTCAATC 439
QY 398 TTGCTGCTCAAGCATACCAAAATCAAGATATGCAAGCAATTTCTTACTTACGATATAT 457
D 440 AATCGGTGACTGATTACCGGAGAGGAAATTTGACACAGCGGTTTATAAAATAGGCGCTAG 499
QY 458 CGCTTCATTTTATGAGAGATGTAATCAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 517
D 500 CAATCCATTTATATACGGATATTAGTCAACTATGCACTATGCACTATGCACTATGCACTATG 559
QY 518 CTTATCCAAATGGGTTTCCATTTCTAAATACGAAAAATTTTGTGATACAAATAAAAAATAACT 577
D 560 CATACCTTCCAGGCTACACCTGTCATATGAAATATAGTAGATACCATTAACACAAAT 619
QY 578 ATATTGTTTTCAGATAGCAATGGATATTGGAATTTGGAAGGAGCAAAACCCAGAAAGATTGGA 637
D 620 ATCAAGCAACGGAAGACATGTTAGCAAAAAGATTTTGTCTCAGATGACGTGAAAGACTGGC 679
QY 638 TTGAAGGAGCAGCGGTAGCAGCTTAACAGAAATTTCTGGCGTTGTGAACGATACGACAA 697
D 680 TCTATGAAATTCGAAAAAGGCGGAAAGCGGACTACCCGAAAAAATAGTCAATGCGAAAACTA 739
QY 698 AAGATTGCTTTGTTAAAAAGCA 717
D 740 AAAAATCATATTAGTAGGA 759
```

RESULT 6

US-10-398-221-2646

; Sequence 2646, Application US/10398221

; Publication No. US20040018514A1

; GENERAL INFORMATION:

; APPLICANT: KUNST, Frederik

; APPLICANT: GLASER, Philippe

; TITLE OF INVENTION: Listeria innocua, genome and applications

; FILE REFERENCE: 344 702 - US

; CURRENT APPLICATION NUMBER: US/10/398,221

; CURRENT FILING DATE: 2003-03-27

; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061

; PRIOR FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: FR 00/12 697

; PRIOR FILING DATE: 2000-10-04

; NUMBER OF SEQ ID NOS: 4025

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2646

; LENGTH: 870

; TYPE: DNA

; ORGANISM: Listeria monocytogenes EGD

US-10-398-221-2646

Query Match 18.2%; Score 154.4; DB 16; Length 870;
Best Local Similarity 53.1%; Pred. No. 3.3e-27;
Matches 329; Conservative 0; Mismatches 291; Indels 0; Gaps 0;

```
QY 98 CACCGATTTTAAAGATGGTCAGCTGAGGATTAAGCATAAATGAGGGGATTAACCTCTCATTTGT 157
D 140 CACATAAACTTAGTTGGTCCGGGATAACCCGACAAATACTGACGTAAATACGCACTATT 199
QY 158 GGAATGTAATCGTGAATTTGACATCACTGCTCGTAATACACGATTTGTAATCCGAATG 217
D 200 GGCCTTTTAAACAGCGGAAAAATACTAGCTAAAGATGTAATCATATGCGAGCTAATT 259
QY 218 AAATGCAATTTAAATGAGTGGCGTCTGATTTAGAAAAATGGTATTTATTTCTGCTGATT 277
D 260 TAAATGAATGAACCTTAAAAAATTCGATAAAACAAATAGCTCAAGGAATATATGATCGGATC 319
QY 278 ACGAATCCCTTATTATGATATAATAGTACATATGCTTCTCACTTTTATGATCCGGATCTG 337
```

Db 320 ATAAAAATCCATATTTATGATAGTACTAGTACATTTTATCTCAATTTTATTAATCTCGATAGAG 379
Qy 338 GAACACATATATCTCTTTGCGAAACATGCAAAAGAACAGCGCGCAAAATATTTTAACC 397
Db 380 ATAATACTTATTTGCGGGTTTGTCTAATGCGAAATAACAGGAGCAAGATATTTCAATC 439
Qy 398 TTGCTGTCACAGCATACCAAAATCAAGATATGCAAGCAAGCATCTTCTACTAGGATAT 457
Db 440 AATCGGTGACTGATTTACCGAGAGGGAATTTGACACAGCGTTTATAAATTAGGCCTAG 499
Qy 458 CGCTTCATTTATTTAGGAGATGGAATCAGCCCAATGATGCGAGCAAACTTTACGAATCTTT 517
Db 500 CAATCCATTTATATACGATATTTAGTCAACCTATGCGACGCCAATATTTTACCGCAAT 559
Qy 518 CTTATCCAAATGGTTTCCATCTAAATACGAAATATTTTGTGATACAATAAAAAATAACT 577
Db 560 CATACCTCCAGGCTACCACTGTGATATGAAATTTACGTAGTACCATTAAACAAAT 619
Qy 578 ATATTGTTTCAGATAGCATGATATTTGGAATTTGGAAGGAGCAAAACCAAGAGATTGA 637
Db 620 ATCAAGCAACGGAACATGTTAGTACAAAGATTTTGTCTCAGATGACGTGAAAGACTGCG 679
Qy 638 TTGAAGGAGCAGCGGTAGCAGCTAAACAAAGATTTATCTGGCGTTGTAACGATACGACAA 697
Db 680 TCTATGAAATGCGAAAGGCGAAAGCGGACTACCCGAAATAGTCAATGCGAAACTA 739
Qy 698 AAGATTGGTTTGTAAAGCA 717
Db 740 AAAATCATATTTTAGTAGGA 759

RESULT 7

US-10-421-654-31
; Sequence 31, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
US-10-421-654-31

Query Match 10.3%; Score 87.4; DB 16; Length 1422;
Best Local Similarity 53.3%; Pred. No. 8.2e-11;
Matches 284; Conservative 0; Mismatches 231; Indels 18; Gaps 4;
Qy 313 TCTCACTTTTATGATCCGGATACGGAACAACATATATTTCCCTTTTTCGAAACATGCAAAA 372
Db 395 TCTCATTTCTATGATCCTGATACGAAAGAAAGAACTATAAAGGAGAGAGAACCAACAGCC 444
Qy 373 GAAAC---AGGCGCAAAATATTTTAACTTGTCTGTCAGCATACCAAAATCAAGATATG 429
Db 445 CTTTCCGAGGGGATAAATATTTTAAATTTAGCAGGAGATTTATTTAAGAAAGAAAGATTGG 504
Qy 430 CAGCAAGCATTTCTTCTACTAGGATTTATCGCTTCAATTTATTTAGGAGATGTTGAATCAGCA 489
Db 505 AAACAGGCTTTCTATTTATTTAGTGTTCGACGAGTGTCTATAAAGTTTTCATAGTCTTT 564
Qy 490 ATGCATGAGCAAACTTTTACGAATCT---TTCTTATCCAATGGGTTTCCATTTCTAAATAC 546
Db 565 ATGCATGCTGCTAATTTTACAGCTGTCGACGAGGTGCTATAAAGTTTTCATAGTCTTT 624
Qy 547 GAAATTTTGTGATACAAATAAATAAATACTATATTTGTTTCAGATAGCAATGGATATTGG 606

Qy 490 ATGCATGAGCAAACTTTTACGAATCT---TTCTTATCCAATGGGTTTCCATTTCTAAATAC 546
Db 565 ATGCATGCTGCTAATTTTACAGCTGTCGACATGAGTGCATAAATAAGTTTTCATAGCCTTT 624
Qy 547 GAAATTTTGTGATACAAATAAATAAATACTATATTTGTTTCAGATAGCAATGGATATTGG 606
Db 625 GAAATTTATGTAACGACAGTTTCAGACACCGTTTGAAGTGAAGGATGATAGGGAACATAT 584
Qy 607 AATTGGAAGGAGCAAAACCCAGAGATTGGATTTGAAGGAGCAGCGGTAGCAGCTAAACAA 666
Db 685 AATTGG-----GTCAATTTCTGATCCGAAGCAGTGGATACATGAAACAGCGAAACTC 738
Qy 667 GATTATCTCTGGCTGTGTAACGATACGACAAAAGATTGGTTTGTATAAAGCAGCCGTATCT 726
Db 739 GCAAAAGCAGAAATTTATGAATTTACTAGTAGTAA-----TATTTAAATCTCAATATAT 792
Qy 727 CAAGAATATGAGATATAAATGCGTGCAGGAAGTAACACCGGTGACAGGAAAGGTTTAAATG 786
Db 793 AAAGGAAACAAGATCTTTTGGCAACAAGAGTTATGCCAGCTGTCAGAGGAGTTTAGAG 852
Qy 787 GAAGCGCAGCGGTTTACAGCTGGTTTATATTCATTTGTGTTTGTATACGTATGT 839
Db 853 AAAGCGCAAGAAACACGCGCGGATTTATTCATTTATGTTTATAAACAATATGT 905

RESULT 8

US-10-421-654-97
; Sequence 97, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421,654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
US-10-421-654-97

Query Match 10.1%; Score 85.8; DB 16; Length 1422;
Best Local Similarity 53.1%; Pred. No. 2e-10;
Matches 283; Conservative 0; Mismatches 232; Indels 18; Gaps 4;
Qy 313 TCTCACTTTTATGATCCGGATACGGAACAACATATATTTCCCTTTTTCGAAACATGCAAAA 372
Db 395 TCTCATTTCTACGATCCCGATACAGAAAGAACTATAAAGGAGAGAGAACCAACAGCT 444
Qy 373 GAAAC---AGGCGCAAAATATTTTAACTTGTCTGTCAGCATACCAAAATCAAGATATG 429
Db 445 CTTTCTCAAGGAGATAAATATTTTAAATTTAGCAGGTGAATCTTTTAAAGAAAGATGATTGG 504
Qy 430 CAGCAAGCATTTCTTCTACTAGGATTTATCGCTTCAATTTATTTAGGAGATGTTGAATCAGCA 489
Db 505 AAACAGGCTTTCTATTTATTTAGTGTTCGACGAGTGTCTATAAAGTTTTCATAGTCTTT 564
Qy 490 ATGCATGAGCAAACTTTTACGAATCT---TTCTTATCCAATGGGTTTCCATTTCTAAATAC 546
Db 565 ATGCATGCTGCTAATTTTACAGCTGTCGACGAGGTGCTATAAAGTTTTCATAGTCTTT 624
Qy 547 GAAATTTTGTGATACAAATAAATAAATACTATATTTGTTTCAGATAGCAATGGATATTGG 606

```

664 CAAGATATTCTCGCGTTGTGAACGATACGACAAAGATTGGTTGTGTAAAGCAGCGCTTA 723
Db GTGGAATCGGGAACATTTACCAATGACGAGATTAAATCTCACTATAATAAGGAAACAAT 804
Qy 724 TCTCAAGAAATATGCAGATAAATCGCGTGGCGAAGTAACACCGGTGACAGGAAAGCGTTTA 783
Db 805 GCTC-----TTTGGCAACAAGAAGTTATGCCAGCTGTCAGAGGAGTTTA 849
Qy 784 ATGGAAGCGCAGCGCGTTTACAGCTGGTTATATTCATTTTGGTGTGANTGATGT 839
Db 850 GAGAACGCACAAAGAAACACGCGCGGATTTATTCATTTATGTTTAAACATTTGT 905

RESULT 10
US-10-421-654-81
; Sequence 81, Application US/10421654
; Publication No. US20040005604A1
; GENERAL INFORMATION:
; APPLICANT: Gramatikova, Svetlana
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Lam, David E.
; APPLICANT: Barton, Nelson R.
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-094001
; CURRENT APPLICATION NUMBER: US/10/421.654
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/374,313
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
US-10-421-654-81

Query Match 8.5%; Score 72; DB 16; Length 1422;
Best Local Similarity 51.7%; Pred. No. 4.5e-07;
Matches 277; Conservative 0; Mismatches 235; Indels 24; Gaps 4

Qy 313 TCTCATTTTATGATCCGGATCTGGAAACAATATATCTCTTTTCGAAAC--ATGCA 369
Db 385 TCTCATTTTTCGATCCCGATACAGAAAGAACTATAAAGGGGAAGAAACCAACAGCT 444
Qy 370 AAAGAAACAGCGCGCAAAATATTTTAACTTGCTGCTCAAGCATACCAAAATCAAGATATG 429
Db 445 CTTTCACAGGAGATAAATATTTTAAATAGCAGTGATCTTTAAGAGGGCGACCA 504
Qy 430 CAGCAAGCATTTCTTACTAGGATTTATCGCTTCATATTTAGGAGATGTGATCAGCA 489
Db 505 AAACAAGCTTTTATTTATTTAGGTGTGCAACGCATTACTTTTACAGATGCTACTCAACCA 564
Qy 490 ATGCATGACGCAAACTTTACGAATCT---TTCTTATCCAAATGGGTTTCATTTCTAAATAC 546
Db 565 ATGCATGCTGCTAAATTTTACAGCGCTGACACAGTGCTTTTAAAGTTTCATAGCGCTTTT 624
Qy 547 GAAAAATTTTGTGTGACATAAAAAATACTATATTTGTTTCAGTAGCAATCGATATTGG 606
Db 625 GAAATATGTGACGCAATTCAGACACAGTATGAATATCTGTATGTTGAGGCGGTATAT 684
Qy 607 AATT---GGAAGGAGCAAAACCCAGNAGATTGGATTGAGGAGCAGCGGTAGCAGCTAAA 663
Db 685 AATTTAGTGAATTTCTAATGATCCAAACAGTGGATCCATGAAACAGCAGAGATCTCGCAAA 744
Qy 664 CAAGATTATCTTGGCGTTGTGAACGATACGACAAAGATTGGTTGTGTAAAGCAGCGCTTA 723
Db 745 GTGGAATCGGGAACATTTACCAATGACGAGATTAAATCTCACTATAATAAGGAAACAAT 804
Qy 724 TCTCAAGAAATATGCAGATAAATCGCGTGGCGAAGTAACACCGGTGACAGGAAAGCGTTTA 783

```

Db	625	GAAGATTATGTGACGACAATTCAGGACACAGTTTAAAGTATCAGATCGACGAGGAAAAATAT	684
Qy	607	AATTTGAAAGGACCAAAACCCAGAAGATTGGATTGAAGGAGCAGCGGTAGCAGCTAAACAA	666
Db	685	AAATT-----TAGTAAATTTCTAATGATCCGAAACAGTGGATCCATGAACACGCGAGACTC	738
Qy	667	GATTATCTCGGCGTTGTGAACGATAGACAAAAGATTGGTTGTAAAGCAGCGCGTATCT	726
Db	739	GCAAAAGTGGAAATCGGAAACATTACCA-----ATGATGTGATTAAATCTCACTATAAT	792
Qy	727	CAAGAATATGACATAAAATGGCGTGGGAAAGTAACACCGGTGACAGAAAGCGTTTAATG	786
Db	793	AAAGGAAACAATGCTCTTTGGCAGCAAGAGTTATGCGACGCTTTCAGAGAAGTTTAGAA	852
Qy	787	GAAGCGCAGCGCTTACAGCTGGTGTATATTCATTTTGGTTTTCATACGTTATGT	839
Db	853	CAAGCCCAAGAAATACGGCGGATTTATTCATTTTATGGTTTAAACATATGT	905
RESULT 9			
US-10-421-654-89			
; Sequence 89, Application US/10421654			
; Publication NO. US20040005604A1			
; GENERAL INFORMATION:			
; APPLICANT: Gramatikova, Svetlana			
; APPLICANT: Hazlewood, Geoff			
; APPLICANT: Lam, David B.			
; APPLICANT: Barton, Nelson R.			
; TITLE OF INVENTION: PHOSPHOLIPASES, NUCLEIC ACIDS ENCODING THEM AND			
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM			
; FILE REFERENCE: 09010-094001			
; CURRENT APPLICATION NUMBER: US/10/421.654			
; CURRENT FILING DATE: 2003-04-21			
; PRIOR APPLICATION NUMBER: US 60/374,313			
; PRIOR FILING DATE: 2002-04-19			
; NUMBER OF SEQ ID NOS: 106			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 89			
; LENGTH: 1422			
; TYPE: DNA			
; ORGANISM: Unknown			
; FEATURE:			
; OTHER INFORMATION: Obtained from an environmental sample			
US-10-421-654-89			
Query Match 9.0%; Score 76.8; DB 16; Length 1422;			
Best Local Similarity 52.2%; Pred. NO. 3.1e-08;			
Matches 280; Conservative 0; Mismatches 232; Indels 24; Gaps 4;			
Qy	313	TCTCACTTTATGATCCGGTACTGGAACAACATATATTCCTTTTCGAAACATGCAGAAA	372
Db	395	TCTCATTTTACGATCCCGATACGAAAGAAGAACTATAAGGGGAGAGAACCAACAGCT	444
Qy	373	GA AAC- --AGGCGCAAAATATTTTAACTTGTGTGCTCAAGCATACCAAAATCAAGATG	429
Db	445	CTCTCTCAAGGAGATAAATATTTTAAATAGCAGCGGATTACTTTTAAAGAAAGAGATTGG	504
Qy	430	CAGCAAGCATCTTCTACTTAGGATATTCGCTTCAATTTTAGGAGATGTGAATCAGCCA	489
Db	505	AAACAAGCTTCTTATATTTAGTGTTGGACGCACTACTTCCACAGATGCTACTCAGCCA	564
Qy	490	ATGCAATGCAGCAAACTTTACGAATCT---TTCCTTATCAATGGGTTTCCATCTTAAATAC	546
Db	565	ATGCATGCTGCTAAATTTTACAGCGTCGACACGAGTGTCTTTAAAGTTTTCATAGCGCTTTT	624
Qy	547	GAAATTTTGTGTATACATAAAAAATACTATATTTGTTTCAGATAGCAATGGATATGG	606
Db	625	GAAATTTATGTGACGCAATTCAGACACAGTATGAAGTATCTGTATGGTGAGGCGGTATAT	684
Qy	607	AAATT---GAAAGGAGCAAAACCCAGAAGATTGGATTGGATTGAAGGAGCAGCGGTAGCAGCTAAA	663
Db	685	AAATTAGTGAATTTCTAATGATCCAAACAGTGGATCCATGAACACGAGACTCGCAAAA	744

Db 805 GCTC-----TTTGGCAACAAGAGTTATGCCAGCTGTCCAGAGGAGTTTA 849

QY 784 ATGGAAGCGCAGCGGCTTACAGCTCGTGTATATTTCAATTTGTGTTGATACGATGT 839

Db 850 GAGAAGCACAAGAAACACGCGGGGATTTATTCATTTATGTTTAAACATTTGT 905

RESULT 11

US-10-398-221-3772/c

Sequence 3772, Application US/10398221

Publication No. US20040018514A1

GENERAL INFORMATION:

APPLICANT: KUNST, Frederik

APPLICANT: GLASER, Philippe

TITLE OF INVENTION: Listeria innocua, genome and applications

FILE REFERENCE: 344 702 - US

CURRENT APPLICATION NUMBER: US/10/398,221

CURRENT FILING DATE: 2003-03-27

PRIOR APPLICATION NUMBER: PCT/FR 01/03 061

PRIOR FILING DATE: 2001-10-04

PRIOR APPLICATION NUMBER: FR 00/12 697

PRIOR FILING DATE: 2000-10-04

NUMBER OF SEQ ID NOS: 4025

SOFTWARE: PatenIn version 3.0

SEQ ID NO 3772

LENGTH: 4562

TYPE: DNA

ORGANISM: Listeria monocytogenes 4b

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(end)

OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u

US-10-398-221-3772

Query Match 6.8%; Score 57.4; DB 16; Length 4562;

Best Local Similarity 53.3%; Pred. No. 0.0029; Mismatches 106; Indels 0; Gaps 0;

Matches 121; Conservative 0;

QY 107 TAAAGATGGTCAGCTGAGGATAGCATATAGCGGGATTAACTCTCATTTGTGGATTGTA 166

Db 227 TTAGTTGGTCCGGGTAATCCGACAAATACCTGAGTAATAGCACTATTTGCTTTT 168

QY 167 ATCGTGAATTTGACATCATGCTCTCGTAATACACGATTTGTAATCGGAATGAACCTGCAT 226

Db 167 AACAGCAGAAAAAATACCTAGCTAAAGATGTAGATCATATGCGAGCTAATTTAATGAATG 108

QY 227 TATTAATGAGTGGCGTCTGATTTAGAAAAAGTGTATTTCTGCTGTGATTACGAGATC 286

Db 107 AACTTAAAAATTCGACAAACAAATTTGCTCAAGGAATATATGACGCGGATCATAAAAATC 48

QY 287 CTATTATGATAATAGTACATATGCTTCTCCTCTTATGATCCGGAT 333

Db 47 CATATTATGATACTAGTAGCTTTTATCTCATTTTATTAATCTCAT 1

RESULT 12

US-08-781-986A-470

Sequence 470, Application US/08781986A

Publication No. US20030054436A1

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 470:

SEQUENCE CHARACTERISTICS:

LENGTH: 9821 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-470

Query Match 5.1%; Score 43.4; DB 8; Length 9821;

Best Local Similarity 44.3%; Pred. No. 11;

Matches 176; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

QY 321 TTATGATCGGATACCTGGGAACAACATATATCTCTTTGCGAAACATGCAAAAGAACAGC 380

Db 4223 TGATGTCAGATTTTCAGAAATGAAAGAAAGAAATTTGAAAGCTTGAANAATCAAGAGC 4282

QY 381 CGCAAAATATTTTAACTTGTCTGTCGTCACAGCATACCAAAATCAAGATATGCGACAAAGCAT 440

Db 4283 TGACATCACTGTTAAAGATTCAGTAAGCTGAAAAGAGCAAGAGCGTATTTTAGTAAG 4342

QY 441 CTCTTACTTAGGATTTATCGCTTCATTTTAGGAGATGTAATCAGCCATCAGCCATCAGCGC 500

Db 4343 AATGCAAGAAACACAGAAATGCTTATTTCAATAGCAGCAGCAAGCAATTTAAAGAGC 4402

QY 501 AAACCTTTACGATCTTTCTTATCCATGCTTTCCATTTCTAATACGAAAAATTTTGTGCA 560

Db 4403 AGAAAAAGCAAGAAAGCAAGAAAAAGAGTGAATAGCATAGCATAGATGATGTCAT 4462

QY 561 TACATAAAAAATACTATATTTGTTTCAGATAGCAATGGATATTTGGAATTTGAAAAGGAGC 620

Db 4463 TCCTATAAAAAATACTACCTCAACCTTTCTAAGTCTGAAAAGATAAATTTGTTAGTATTGC 4522

QY 621 AAACCCAGAAATTTGGATTGAAGGAGCAGCGGTAGCAGCTAAACAGATTTATCCTGGCGT 680

Db 4523 TGATCAAGACATAGGATGAAGTAAGAAAGGCAAAATCTAAAAAAGATGCTGTAGTAGA 4582

QY 681 TGTGAACGATACGACAAAGATTTGTTGTAAGCA 717

Db 4583 CGTTGTTAAAAAGCAAAATAAGATTTGTAAGAA 4619

RESULT 13

US-10-329-624-470

Sequence 470, Application US/10329624

Publication No. US20040040307A1

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

```

; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248PID1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
;
; INFORMATION FOR SEQ ID NO: 470:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9821 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 470:
US-10-329-624-470

Query Match 5.1%; Score 43.4; DB 13; Length 9821;
Best Local Similarity 44.3%; Pred. No. 11;
Matches 176; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

QY 321 TTATGATCGGATCTGGACACATATATCTCTTGGAAACATGCAGAAACAGG 380
Db 4223 TGATGTCAGATTTCAGAAATGAAAGAAAGAAATTTGAAAGCTTGAATCAAGAGC 4282

QY 381 CGCAAAATATTTTAACTTGGTCAAGCATACCAAAATCAAGATATGCAGCAAGCAT 440
Db 4283 TGACATCACTGTTAAGAATTGAGTAAGCTGAAAGAGCAGAGCGTATTTTAGTAAG 4342

QY 441 CTCTACTTAGGATTATCGCTTCATTTATTTAGGAGATGGAATCAGCCAAATCGATGCGC 500
Db 4343 AATGCAAGAAACAGAAATGCTTATTTCAATAGACGAGCGCAAGCAATTTAAAGAGC 4402

QY 501 AAACCTTAGCAATCTTTCTTATCCATGGTTTCCATTTCTAATACGAAATTTTGTGA 560
Db 4403 AGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGC 4462

QY 561 TACAATAAAAATAACTATATTTTTCAGATAGCAATGATATTTGAAATTTGAAAGGAGC 620
Db 4463 TGTATATAAATAATAGCTCAACCTTTCTAAGTCTGAAAGAGATAAATTTAGCTATTGC 4522

QY 621 AAACCCAGAGATTGGATTGAAGGAGCAGCGGTAGCAGCTAAACAGAGATTATCTGGCGT 680
Db 4523 TGATCAAGACATAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4582

QY 681 TGTGAACGATACGACAAAGATTGGTTTCTTAAAGCA 717
Db 4583 CGTTGTTAAAGCAAAATAAGATATTGATAAGAA 4619

RESULT 14
US-09-966-521-21/c
; Sequence 21, Application US/09966521
; Publication No. US20030087321A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 00774.US1 CN1
; CURRENT APPLICATION NUMBER: US/10/429,094
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 09/966,521
; PRIOR FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-429-094-21

Query Match 5.1%; Score 43.2; DB 15; Length 837;
Best Local Similarity 50.5%; Pred. No. 3.4;
Matches 105; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 403 GGTCAAGCATACCAAAATCAAGATATGCAGCAAGCATTTCTTACTTAGGATTATCGCTT 462
Db 820 GGCCAAATCGACCTTGCACAGTATGTCAGAAATACTGGTGTAGTTACATTATATCGT 761

QY 463 CATTATTTAGGAGATGTAATCAGCCAAATGCATGCAAGCAAACTTTACGAAATCTTTCTAT 522
Db 760 AATTATTCACATGAAAGCTTCATTAAAACTTTCTTCAATTTATCAACATATTCGAATGA 701

QY 523 CCAATGGTTTCCATTTCTAATACGAAATTTTGTGTATACATAAATAAATACTATAT 582
Db 700 CGTTAGCATGTGCGACACCAATGGAATTTGATGATCTCTTATAAATTCAGCAATTT 641

QY 583 GTTTCAGATAGCAATGGATATTGGAATT 610
Db 640 CTTTTCAGTATTTGATAGTACTAGAAATT 613

RESULT 15
US-10-429-094-21/c
; Sequence 21, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 00774.US1 CN1
; CURRENT APPLICATION NUMBER: US/10/429,094
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 09/966,521
; PRIOR FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-429-094-21

Query Match 5.1%; Score 43.2; DB 15; Length 837;
Best Local Similarity 50.5%; Pred. No. 3.4;
Matches 105; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 403 GGTCAAGCATACCAAAATCAAGATATGCAGCAAGCATTTCTTACTTAGGATTATCGCTT 462
Db 820 GGCCAAATCGACCTTGCACAGTATGTCAGAAATACTGGTGTAGTTACATTATATCGT 761

QY 463 CATTATTTAGGAGATGTAATCAGCCAAATGCATGCAAGCAAACTTTACGAAATCTTTCTAT 522
Db 760 AATTATTCACATGAAAGCTTCATTAAAACTTTCTTCAATTTATCAACATATTCGAATGA 701
```

Qy	523	CCAATGGGTTCCATTCTTAATAACGAAATTTTGTGTGATACAAATAAAAAATAACTATATT	582
Db	700	CGTTAGCATGTGCGACACCAATGGATTGATTTTCATGATCTCCTATAAAATTCAGCAATT	641
Qy	583	GTTCAGATAGCAATGGATATTGGAATT	610
Db	640	CCTTTTCAAGTATTGGATACTAGAATT	613

Search completed: August 2, 2004, 04:20:53
 Job time : 511 secs

This page blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2004, 23:39:34 ; Search time 432 Seconds
(without alignments)
8348.889 Million cell updates/sec

Title: PCT-US03-12556-1
Perfect score: 849
Sequence: 1 atgaaaagaagattagc.....atcgtatgtaaatcgctaa 849

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001as:.*
5: Geneseq2001bs:.*
6: Geneseq2002s:.*
7: Geneseq2003as:.*
8: Geneseq2003bs:.*
9: Geneseq2003cs:.*
10: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	154.4	18.2	870	6	ABQ67927 Listeria
2	154.4	18.2	870	6	ABQ69833 Listeria
3	154.4	18.2	110000	6	Continuation (3 of
C 4	57.4	6.8	4562	6	ABQ70959 Listeria
C 5	46.6	5.5	3706	2	AAT36392 Lactococc
6	45.8	5.4	9821	2	AAV74781 Staphyloc
7	45.4	5.3	6033	3	AAA70152 Plasmodi
8	45.2	5.3	3738	3	AAA70178 Plasmodi
9	45	5.3	2000	7	ADA71938 Rice gene
10	43.4	5.1	6222	7	ACF72350 Staphyloc
C 11	43.2	5.1	837	7	ACF74582 Staphyloc
C 12	43.2	5.1	837	9	ADC21339 Staphyloc
C 13	43.2	5.1	837	9	ADC25025 DNA encod
C 14	43.2	5.1	837	9	ADD52527 Staphyloc
C 15	43.2	5.1	2700	3	AAA70220 Plasmodi
C 16	43.2	5.1	5924	2	AAV74441 Staphyloc
C 17	43	5.1	8170	6	ABK28258 DNA trans
18	43	5.1	43095	3	AAA68254 Bacteriop
C 19	42	4.9	969	6	ABQ39490 Oligonucle
C 20	42	4.9	969	6	ABQ39491 Human imm
C 21	41.8	4.9	6227	6	ABL33586 Chemical
C 22	41.8	4.9	6227	6	ABL92260 Chemical
C 23	41.6	4.9	2319	4	AAC84742 S. pneumo

24	41.6	4.9	2911	2	AAV52332 Streptoco
C 25	41.6	4.9	3840	7	ABQ77399 Human col
26	41.6	4.9	6361	3	AAZ91806 Streptoco
27	41.6	4.9	6420	7	ABX06302 S. pneumo
28	41.6	4.9	6423	7	ACA49616 Prokaryot
C 29	41.6	4.9	110000	6	Continuation (17 o
30	41.6	4.9	110000	6	ABQ67195 Listeria
31	41.6	4.9	110000	7	Continuation (5 of
32	41.6	4.9	110000	7	Continuation (6 of
C 33	41.2	4.9	2000	7	ADA71938 Rice gene
34	41	4.8	3491	4	ABL19460 Drosophil
C 35	41	4.8	4985	6	ABQ75107 Anopheles
C 36	41	4.8	4985	9	ACF79720 Mosquito
37	40.8	4.8	3837	3	AAA70211 Plasmodi
C 38	40.2	4.7	1014	7	ACA43139 Prokaryot
39	40	4.7	3381	4	AAZ51736 Staphyloc
40	40	4.7	3477	4	AAZ54787 Staphyloc
41	39.8	4.7	1113	2	AAQ62305 Clostridi
C 42	39.8	4.7	1353	2	AAZ99537 Nucleic a
43	39.8	4.7	2130	7	ACA52696 Prokaryot
44	39.4	4.6	705	2	AAQ10244 Streptoco
45	39.4	4.6	1617	2	AAT49317 Type-6 M-

ALIGNMENTS

RESULT 1
ABQ67927
ID ABQ67927 standard; DNA; 870 BP.
XX
AC ABQ67927;
XX
DT 29-AUG-2003 (revised)
DT 29-AUG-2002 (first entry)
XX
DE Listeria monocytogenes EGD DNA sequence #51.
XX
KW Antibacterial; Listeria; food contamination; mutational analysis;
KW infection; ds.
XX
OS Listeria monocytogenes; EGD.
XX
PN WO200228891-A2;
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-FR003061.
XX
PR 04-OCT-2000; 2000FR-00012697.
XX
PA (INSP) INST PASTEUR.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Kunst F, Glaser P;
XX
DR WPI; 2002-332479/37.
XX
PT New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators.
XX
PS Claim 16; SEQ ID NO 740; 180pp; French.
XX
CC The present invention relates to nucleic acid sequences (ABQ67188-
CC ABQ71212) from Listeria sp. The sequences are useful as probes and
CC primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of gene
CC expression. Proteins encoded by the nucleic acid sequences can be used to
CC screen for compounds that modulate gene expression, replication and
CC pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in anti-
CC Listeria vaccines. Note: The sequence data for this patent did not form


```
Qy 638 TTGAAGGAGCGGTAGCAGCTAAACAAGATTATCTCTGGCGTGTGTAACGATACGACAA 697
Db 680 TCTATGAAATGCGAAGGCGAAGCGGACTACCCGAAATAGTCAATGCGAAACTA 739

Qy 698 AAGATTGGTCTTGTAAAGCA 717
Db 740 AAAAATCATATTTAGTAGGA 759

RESULT 3
ABAO3041_02
Continuation (3 of 30) of ABA03041 from base 200001 (Listeria monocytogenes EGD-e genome
WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041
WP Fragment Name Begin End
WP ABA03041_00 1 110000
WP ABA03041_01 100001 210000
WP ABA03041_02 200001 310000
WP ABA03041_03 300001 410000
WP ABA03041_04 400001 510000
WP ABA03041_05 500001 610000
WP ABA03041_06 600001 710000
WP ABA03041_07 700001 810000
WP ABA03041_08 800001 910000
WP ABA03041_09 900001 1010000
WP ABA03041_10 1000001 1110000
WP ABA03041_11 1100001 1210000
WP ABA03041_12 1200001 1310000
WP ABA03041_13 1300001 1410000
WP ABA03041_14 1400001 1510000
WP ABA03041_15 1500001 1610000
WP ABA03041_16 1600001 1710000
WP ABA03041_17 1700001 1810000
WP ABA03041_18 1800001 1910000
WP ABA03041_19 1900001 2010000
WP ABA03041_20 2000001 2110000
WP ABA03041_21 2100001 2210000
WP ABA03041_22 2200001 2310000
WP ABA03041_23 2300001 2410000
WP ABA03041_24 2400001 2510000
WP ABA03041_25 2500001 2610000
WP ABA03041_26 2600001 2710000
WP ABA03041_27 2700001 2810000
WP ABA03041_28 2800001 2910000
WP ABA03041_29 2900001 2944528

Query Match 18.2%; Score 154.4; DB 6; Length 110000;
Best Local Similarity 53.1%; Pred. No. 1.6e-28;
Matches 329; Conservative 0; Mismatches 291; Indels 0; Gaps 0;

Qy 98 CACCGATTTTAAAGATGGTCAGCTGAGGATAAGCATATGAGGGGATTAACTCTCATTTGT 157
Db 11564 CACATAAATCTAGTTGGTCCGGGATACCCGACAAATACTGACGTAATAGGCACTATT 11623

Qy 158 GGATTTGTAATCGTCAATGACATCATGTCTCGTAATACACGATGTGTAATCGCAATG 217
Db 11624 GGCTTTTAAACAGCGGAAAAAATACTAGCTAAAGATGTAATCATATGCGAGCTAAT 11683

Qy 218 AAATGCTATTATTAATGATGGCGTGTGATTAGAAATGGTATTATTTCTGCTGATT 277
Db 11684 TAATGAATGAACCTTAAAAAATTCGATAAACAATAGCTCAAGGAATATATGATCGGATC 11743

Qy 278 ACGAGAACCTTATTATGATAATAGTACATATGCTTCTCACCTTTTATGATCGGATACG 337
Db 11744 ATAAAAATCCATATTATGATAGTACTAGTACATTTTATCTCATTTTATATCTCTGATAG 11803

Qy 338 GAACAACATATATTCCTTTTGGAAACATGCAAAAGAACAGCGCCAAATAATTTAAACC 397
Db 11804 ATAATACTTATTTGCGGGTTTGTCTAATGCGAAAAATAACAGGACAAAGATTATTTCAATC 11863

Qy 398 TTGCTGGTCAAGCATACCAAAATCAAGATATGACCAAGCATCTTCTTACTTAGGATAT 457
Db 11864 AATCGGTGACTGATTACCGGAGAGGAAATTTGACACGCGTTTATAAATTAGGCCTAG 11923
```

```
Qy 458 CGCTTATTATTATTAGGAGATGTGAATCAGCAATGCAATGCAAGCAAACTTTTACGAATCTTT 517
Db 11924 CAATCCATTATTATACGGATATTAGTCAACTATGCAAGCAATTAATTTTACCGCAATAT 11983

Qy 518 CTTATCCAAATGGGTTTCCATTCTTAATACGAAATTTTGTGATACATATAAATAACT 577
Db 11984 CATACCTCCAGGCTACCACTGTGCATATGAATTTACGTAGTACCACTTAAACCAAT 12043

Qy 578 ATATTGTTTTCAGATAGCAATGATATTGGAATTTGAAAGGAGCAAAACCCAGAAAGATTGA 637
Db 12044 ATCAAGCAACGAGACATGTTAGCAAAAGATTTTGTCTCAGATGACGTGAAGACTGGC 12103

Qy 638 TTGAAGGAGCGGTAGCAGCTAAACAAGATTATCTCTGGCGTGTGTAACGATACGACAA 697
Db 12104 TCTATGAAATGCGAAGGCGAAGCGGACTACCCGAAATAGTCAATGCGAAACTA 12163

Qy 698 AAGATTGGTCTTGTAAAGCA 717
Db 12164 AAAAATCATATTTAGTAGGA 12183

RESULT 4
ABQ70959/C
ID ABQ70959 standard; DNA; 4562 BP.
XX
AC ABQ70959;
XX
DT 29-AUG-2003 (revised)
DT 29-AUG-2002 (first entry)
XX
DE Listeria monocytogenes 4b contig DNA sequence #901.
XX
KW Antibacterial; Listeria; food contamination; mutational analysis;
KW infection; ds.
XX
OS Listeria monocytogenes ATCC 19115.
XX
WO200228891-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-FR003061.
XX
PR 04-OCT-2000; 2000FR-00012697.
XX
PA (INSP ) INST PASTEUR.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Kunst F, Glaser P;
XX
WPI; 2002-332479/37.
XX
PT New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators.
XX
PS Claim 14; SEQ ID NO 3772; 180pp; French.
XX
CC The present invention relates to nucleic acid sequences (ABQ67188-
CC ABQ71212) from Listeria sp. The sequences are useful as probes and
CC primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of gene
CC expression. Proteins encoded by the nucleic acid sequences can be used to
CC screen for compounds that modulate gene expression, replication and
CC pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in anti-
CC Listeria vaccines. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. (Updated
CC on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 4562 BP; 1166 A; 943 C; 782 G; 1671 T; 0 U; 0 Other;
```

Query Match 6.8%; Score 57.4; DB 6; Length 4562;
 Best Local Similarity 53.3%; Pred. No. 0.00024;
 Matches 121; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy 107 TAAGATGCTGAGCTGAGGATAAGCATTAATGAGGGGATTAACCTCTCATTTGTGGATTCTAA 166
 Db 227 TTAGTTGGTCCGGGATAATCCGCAATACTAGCAATAGTACGTAATACGCACTATTGGCTTTTA 168

Qy 167 ATCGTGAATTTGACATCATGCTCTGTAATACACGATTTGGAATCCGAATGAATCTGCAT 226
 Db 167 AACAGCAGAAAAAATACTAGCTAAAGATGTAGATCATATGCGAGCTAAATTTAATGAATG 108

Qy 227 TATTAAATGAGTGGCGTCTGATTTAGAAATGATTTATTTCTCTGATTTACGAGATC 286
 Db 107 AACTTAAAAATTTTCGACAAACAAATGCTCAAGGAATATATGACGCGGATCAATAAAATC 48

Qy 287 CTTATTATGATAAGTACATATGCTCTCACTTTTATGATCCGAT 333
 Db 47 CATATTATGATACTAGTACGTTTATCTCATTTTATATCTCTGAT 1

RESULT 5
 AAT36392/C
 ID AAT36392 standard; DNA; 3706 BP.

XX AAT36392;
 AC
 XX
 DT 16-OCT-2003 (revised)
 DT 08-DEC-1996 (first entry)

XX Lactococcus lactis subsp. cremoris W9 restriction-modification system.
 XX
 KW Lactic acid bacterium; Danish starter culture; cheese; enzyme;
 KW restriction endonuclease; methylase; fermented milk; phage resistance;
 KW ds.

OS Lactococcus lactis subsp. cremoris; W9.
 XX
 FH Key Location/Qualifiers
 CDS complement(422..2161)
 FT /tag= a
 FT /codon_start= 422
 FT /product= "LlaBI methylase"
 FT /evidence= EXPERIMENTAL
 FT /gene= "ORF"
 FT /number= 1 /standard name= "Gene coding for LlaBI
 FT methylase
 FT /label= m-llaBI
 FT 2464..3360
 FT /tag= b
 FT /codon_start= 2464
 FT /product= "LlaBI restriction endonuclease"
 FT /evidence= EXPERIMENTAL
 FT /gene= "ORF"
 FT /number= 2 /standard name= "Gene coding for LlaBI
 FT endonuclease
 FT /label= r-llaBI

XX WO9625503-A1.
 XX
 XX 22-AUG-1996.
 XX
 XX 19-FEB-1996; 96WO-DK000076.
 XX
 XX 17-FEB-1995; 95DK-00000179.
 XX
 XX (JOSE/) JOSEPHSEN J.
 XX (NYEN/) NYENGAARD N. R.
 XX (VOGE/) VOGENSEN F. K.
 XX (MADS/) MADSEN A.
 XX Josephsen J, Nyengaard NR, Vogensen FK, Madsen A;

XX WPI; 1996-393404/39.
 DR P-PSDB; AAW02164, AAW02165.
 XX
 PT Type II restriction-modification systems obtd. from Lactococcus lactis -
 PT for conferring phage resistance on lactic acid bacteria, useful as
 PT starter cultures for cheese and fermented milk prods.
 XX
 XX Claim 2; Page 67-69; 93pp; English.

XX This DNA encodes the plasmid-derived type-II restriction-modification
 CC system from Lactococcus lactis subsp. cremoris W9, llaBI, and contains 2
 CC ORFs transcribed in the same direction and coding for putative proteins
 CC of 580 AA (llaBI methylase M.LlaBI; AAW02164) and 299 AA (llaBI
 CC endonuclease R.LlaBI; AAW02165). This restriction-modification system may
 CC be used in a method for conferring increased virus resistance, more
 CC specifically phage resistance, to a L. lactis strain used in cheese
 CC manufacture. (Updated on 16-OCT-2003 to standardise OS field)

XX
 SQ Sequence 3706 BP; 1284 A; 535 C; 575 G; 1312 T; 0 U; 0 Other;

Query Match 5.5%; Score 46.6; DB 2; Length 3706;
 Best Local Similarity 47.7%; Pred. No. 0.12; Indels 0; Gaps 0;
 Matches 136; Conservative 0; Mismatches 149;

Qy 51 AGTTCAAAGTGTAGTATTTCACAAACAAATAATAGTGAAGTCTCTCCCGATTTTAAAG 110
 Db 1010 ATTATATCGAATAATATATCTTATGAATAATATGACGCAATGCTGTACCAATCAGCAT 951

Qy 111 ATGGTCAGCTGAGGATAAGCATTAATGAGGGGATTAACCTCTCATTTGGATTGTAATCG 170
 Db 950 TGTCTCAATGAAGGAAAAATTCCTTAATACCTTTAAATTTATTTATTCCTATTAAAGACGA 891

Qy 171 TGCATTTGACATCATGCTCTGTAATACAGGATTTGTGAATCCGAATGAACTGCATTATT 230
 Db 890 GCTCAGTAAACGCAATGCTGTGGAACCAACCGTTAACATTTGGTATGATATGGCGATC 831

Qy 231 AAATGAGTGGCGTCTGATTTAGAAAAATGGTATTATTCTCTGCTGATTACGGAATCTTTA 290
 Db 830 TCAAGCATTTGAATAGCCATGTACCTAAATAATTTTCTTCAAAATTTCTTTGAATCCAAA 771

Qy 291 TTATGATAATAGTACATATGCTTCTCACTTTTATGATCCGGATAC 335
 Db 770 CTTTGTATATTTTACAGATTATGCTCTTTTAAATAATGGATATGC 726

RESULT 6
 AAV74781
 ID AAV74781 standard; DNA; 9821 BP.

XX AAV74781;
 AC
 XX
 XX 16-MAR-1999 (first entry)
 DT
 XX
 DE Staphylococcus aureus contig SEQ ID #470.
 XX
 KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.

XX Staphylococcus aureus.
 OS
 XX
 XX Key Location/Qualifiers
 FH misc_feature 1021..1080
 FT /tag= a
 FT /note= "these bases represent a line of missing text in
 FT the sequence listing in the specification. they are
 FT included to maintain the nucleotide numbering given in
 FT the specification for this DNA sequence"
 FT 2821..2880
 FT misc_feature
 FT /tag= b
 FT /note= "these bases represent a line of missing text in
 FT

the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

4621..4680

/*tag= c

/note= "these bases represent a line of missing text in the sequence listing in the specification. they are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

6421..6480

/*tag= d

/note= "these bases represent a line of missing text in the sequence listing in the specification. they are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

8221..8280

/*tag= e

/note= "these bases represent a line of missing text in the sequence listing in the specification. they are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

EP786519-A2.

30-JUL-1997.

07-JAN-1997; 97EP-00100117.

05-JAN-1996; 96US-0009861P.

(HUMA-) HUMAN GENOME SCI INC.

Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

WPI; 1997-374922/35.

Polynucleotide(s) and proteins derived from *Staphylococcus aureus* - stored on computer readable medium and used in the production of anti-S. aureus vaccines.

Claim 1; Page 1387-1392; 3271pp; English.

This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences of the invention. The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the S. aureus DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against S. aureus infection. The polypeptides can also be used in a kit for the immunodetection of S. aureus in a sample. S. aureus is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the S. aureus DNA sequences contained on the computer readable medium

Query Match 5.4%; Score 45.8; DB 2; Length 9821;

Best Local Similarity 38.8%; Pred. No. 0.26;

Matches 179; Conservative 0; Mismatches 282; Indels 0; Gaps 0;

Qy 321 TTATGATCCGGATCTGGAACAACATATATTCCTTTTCGAAACATGCAAAAGAAACAGG 380

|||||

Db 4223 TGATGGTCAGATTTGAGAAATGAGAAAGAAATGCAAAAGCTTGAATCAAGACG 4282

|||||

Qy 381 CGCAAAATATTTTAACTTCTGGTCAAGCATACCAAAATCAAGATATGCAGCAAGCAT 440

|||||

Db 4283 TGACATCACTGTTAAAGAAATTGAGTAAGACTGAAAGAGCAAGAGCGTATTTTAGTAG 4342

Qy 441 CTTCCTACTTACGATTTATCGCTTTCAATTTATTAGGAGATGTGAATCAGCAATGCATGCAGC 500

|||||

Db 4343 AATGCAAGAAACAGAAATGCTTATTCAATAGACGAGGCAAGCAATTAAGAAAGC 4402

|||||

Qy 501 AAACCTTTAGGAATCTTTCTTATCCAAATGGGTTTCAATTCCTAAATACGAAAAATTTTGTGA 560

|||||

Db 4403 AGAAAAACCAAGAAAGCAAGAAAAAGAAAGTGGATAAGCAGTATGAAGATGATGTCTAT 4462

|||||

Qy 561 TACAATAAAAAATTAACATATATTGTTTTCAGATAGCAATGATGATATTCGAATTTGAAAGGAGC 620

|||||

Db 4463 TGTATATAAAAAATTAACGTCACCTTTCTAAGTCTGAAAAAGATAAATTTGTAGCTATTGC 4522

|||||

Qy 621 AAACCCAGAGATTGGATTGAAGGAGCAGCGGTAGCAGCTAAACAAAGATTATCTCGGCGT 680

|||||

Db 4523 TGATCAAGACATTAAGGATGAAGTAAGAAAGCAAAATCTTAAAAAGATGCTGTAGTAGA 4582

|||||

Qy 681 TGTGAACGATACGACAAAGATTGGTTTGTAAAGCAGCGGTATCTCAAGATATATGCAGA 740

|||||

Db 4583 CGTTGTTTAAAAAGCAAAATAAAGATATTTGATAAAGAAAAANNNNNNNNNNNNNNNNN 4642

|||||

Qy 741 TAAATGGCGTGGGAAGTAACACCGGTGACAGGAAAGCGTT 781

|||||

Db 4643 NNN 4683

RESULT 7

AAA70152

ID AAA70152 standard; DNA; 6033 BP.

XX

AC AAA70152;

XX

DT 07-NOV-2000 (first entry)

XX

DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:285.

XX

KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;

KW antimalarial; malaria; protozoacide; infection; insecticide; ds.

XX

OS Plasmodium falciparum.

XX

FN WO200025728-A2.

XX

PD 11-MAY-2000.

XX

PF 05-NOV-1999; 99WO-US026796.

XX

PR 05-NOV-1998; 98US-0107131P.

XX

PA (HOFF//) HOFFMAN S.

PA (CARU//) CARUCCI D.

PA (GARD//) GARDNER M.

XX (VENT//) VENTER J C.

PI Hoffman S, Carucci D, Gardner M, Venter JC;

DR WPI; 2000-365347/31.

XX

PT Proteins encoded by chromosome 2 of the human malarial parasite,

PT Plasmodium falciparum, useful as antimalarial vaccines and in the

PT diagnosis of P.falciparum infection.

XX

PS Disclosure; Page 493-495; 577pp; English.

XX

CC The present invention describes proteins and their fragments (I) encoded

CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.

CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)

CC vaccines against P. falciparum infection comprising (I) or (II). (I) and

CC (II) are useful for the development of vaccines against P. falciparum

CC infection. (1) and polyclonal antisera or a monoclonal antibody raised to

CC immunogens comprising the sequences of (I), are useful in the detection

CC of infection with P. falciparum. Furthermore, (I) (especially when they

are rifins or secreted or membrane proteins) can aid the identification of drugs to treat or prevent *P. falciparum* infection, or they can be used to identify drug resistance in *P. falciparum*. Sequencing of the Plasmodium chromosome 2 and the subsequent identification of proteins encoded by it will help to expand our understanding of parasite biology, a process hampered by the complexity of the parasitic lifecycle, and provide new targets for vaccine and drug development. Parasite resistance to drugs and mosquito resistance to insecticides have led to a resurgence of malaria in many parts of the world, and there is a pressing need for vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide and protein sequences given in the present invention, but which are not specifically mentioned within the specification

XX Sequence 6033 BP; 3019 A; 437 C; 707 G; 1870 T; 0 U; 0 Other;

Query Match 5.3%; Score 45.4; DB 3; Length 6033;
Best Local Similarity 45.2%; Pred. No. 0.29;
Matches 166; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 257 ATGGTATTCTGCTGATTACGAGATCCCTTATTGATTAATAGTACATATGCTTCTC 316

DB 5042 ATAAATATTTTATGATATCATGATACCAATAATTAATTAATTAATAATAATA 5101

QY 317 ACTTTATGATCGGATACGAAACACATATATTCCTTTTCGAAACATGCAAAAGAAA 376

DB 5102 ATAATAATAGTAATAATAATAATAATAATATATGCTTAAGAAATAATAATAATA 5161

QY 377 CAGGCGCAAAATATTTTACCTTGCTGTCAGCATACCAAAATCAAGATATGCGAAG 436

DB 5162 TAATGATGAGATATCAATGCAAAATAAAGAGAAAGCTTAAAAAATAAAAAA 5221

QY 437 CATCTCTTCTAGGATATCGCTTCATTATTTAGGAGATGCTGAATCAGCCAAATGATG 496

DB 5222 AAAAAAATGTATACAAAAAATAATAATATTTTGGAAAGAAATAATCAATATTC 5281

QY 497 CAGCAAACTTTACGAACTTTCTTATCCAAATGGGTTTCATTCCTAAATACGAAATTTTG 556

DB 5282 ATAAATAATAGTTTCAAGATACATATTTTAACTGTGAGATCTTTTAAATGAAAGATATG 5341

QY 557 TTGATACAAATAAAAAATACTATATTTTTCAGATAGCAATGATATTTGGAAATGGAAG 616

DB 5342 CTAAATAATAACAAATAAAGTGTGATGAATAATAGCATATCATGTATAATAACATGA 5401

QY 617 GAGCAAA 623

DB 5402 GAGAAAA 5408

RESULT 8

ADA70178

ID AAA70178 standard; DNA; 3738 BP.

XX AAA70178;

XX 07-NOV-2000 (first entry)

XX Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:311.

XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine; antimalarial; malaria; protozoacide; infection; insecticide; ds.

XX Plasmodium falciparum.

XX WO200025728-A2.

XX 11-MAY-2000.

XX 05-NOV-1999; 99WO-US026796.

XX 05-NOV-1998; 98US-0107131P.

XX (HOFF/) HOFFMAN S.

PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
XX (VENT/) VENTER J C.

XX Hoffman S, Carucci D, Gardner M, Venter JC;

XX WPI; 2000-365347/31.

XX Proteins encoded by chromosome 2 of the human malarial parasite,
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of *P. falciparum* infection.

XX Disclosure; Page 511; 577pp; English.

XX The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against *P. falciparum* infection comprising (I) or (II); (I) and
CC (II) are useful for the development of vaccines against *P. falciparum*
CC infection. (I) and polyclonal antisera or a monoclonal antibody raised to
CC immunogens comprising the sequences of (I), are useful in the detection
CC of infection with *P. falciparum*. Furthermore, (I) (especially when they
CC are rifins or secreted or membrane proteins) can aid the identification
CC of drugs to treat or prevent *P. falciparum* infection, or they can be used
CC to identify drug resistance in *P. falciparum*. Sequencing of the
CC Plasmodium chromosome 2 and the subsequent identification of proteins
CC encoded by it will help to expand our understanding of parasite biology,
CC a process hampered by the complexity of the parasitic lifecycle, and
CC provide new targets for vaccine and drug development. Parasite resistance
CC to drugs and mosquito resistance to insecticides have led to a resurgence
CC of malaria in many parts of the world, and there is a pressing need for
CC vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
CC represent nucleotide and protein sequences given in the present
CC invention, but which are not specifically mentioned within the
CC specification

XX Sequence 3738 BP; 1705 A; 258 C; 542 G; 1233 T; 0 U; 0 Other;

Query Match 5.3%; Score 45.2; DB 3; Length 3738;
Best Local Similarity 51.5%; Pred. No. 0.28;

Matches 104; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 192 TAATACACGATTTGTAATCCGATGAACTGCAATTAATAAGTGGCGTCTGATTT 251

DB 2730 TAAGACAATGAAAAATAACCGAACTTAACAAATAAATAATATGATATGACATAT 2789

QY 252 AGAAATGGTATTTATTTCTGCTGATTAAGAGATCCTTATATGATATAGTACATATGC 311

DB 2790 TGATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATG 2849

QY 312 TTCTCACTTTTATGATCCGATATCTGGAACAACATATATTCCTTTTGGAAACATGCAA 371

DB 2850 TGATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 2909

QY 372 AGAAACAGCGCAAAATATTTT 393

DB 2910 AGACTTATATGTAACACATTTT 2931

RESULT 9

ADA71938

ID ADA71938 standard; DNA; 2000 BP.

XX ADA71938;

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice; gene; ds.

XX Oryza sativa.


```
RESULT 11
ACF74582/c
ID ACF74582 standard; DNA; 837 BP.
XX
XX
AC ACF74582;
XX
XX
DT 20-NOV-2003 (first entry)
XX
XX
DE Staphylococcus aureus DNA #2262.
XX
XX
KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW enzymatic assay; antibiotic target; gene; ds.
XX
XX
OS Staphylococcus aureus.
XX
XX
PN WO200294868-A2.
XX
XX
PD 28-NOV-2002.
XX
XX
PF 27-MAR-2002; 2002WO-IB002637.
XX
XX
PR 27-MAR-2001; 2001GB-00007661.
XX
XX
PA (CHIR-) CHIRON SPA.
XX
XX
PI Masignani V, Mora M, Scarselli M;
XX
XX
DR WPI; 2003-120786/11.
DR P-PSDB; ABM73022.
XX
XX
PT New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
PT S. aureus, e.g. sepsis.
XX
XX
PS Claim 6; SEQ ID NO 4523; 49pp; English.
XX
XX
CC The invention relates to novel genes and encoded proteins from
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to Staphylococcus bacteria, specifically an
CC infection caused by S. aureus. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel S. aureus genes of the invention
XX
SQ Sequence 837 BP; 314 A; 116 C; 149 G; 258 T; 0 U; 0 Other;

Query Match 5.1%; Score 43.2; DB 7; Length 837;
Best Local Similarity 50.5%; Pred. No. 0.58;
Matches 105; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 403 GGTCAAGCATACAAATCAAGATATGACGAGCAATCTTCTACTTAGGATTATCGCTT 462
Db 820 GGCCAAATCGACCTTGACCATGATGTCGAGAAATACCTGGTGTAGTACATTATATCGT 761
QY 463 CATTTATTAGGAGATGTAATCAGCAATGATGAGCAAACTTTACGAATCTTCTTAT 522
Db 760 AATTATTACATGAAAGCTTCATTTAAATCTTCTTCAATTTATCAATATCAATGA 701
QY 523 CCAATGGGTTTCCATTTCTAAATACGAAATTTTGTGTATACAAATAAAATACTATTT 582
Db 700 CGTTAGCATGCGGACCAATGATTTGATTTCAATCTTCTTATATTAATTCAGCAATTT 641
QY 583 GTTTCAGATAGCAATGGATTTGGAATT 610
Db 640 CCTTTTCAAGTATTGGATAGTAATT 613

RESULT 12
ADC21339/c
ID ADC21339 standard; DNA; 837 BP.
XX
XX
AC ADC21339;
XX
XX
DT 18-DEC-2003 (first entry)
XX
XX
DE Staphylococcus aureus protein coding sequence SEQ ID NO: 21.
XX
XX
KW ds; gene; antibacterial; gene therapy; vaccine; diagnosis;
KW microbial survival; gene function; bacterial infection;
KW antibiotic-resistant bacterium.
XX
XX
OS Staphylococcus aureus.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..837
FT /tag= a
XX
XX
PN WO2003029484-A2.
XX
XX
PD 10-APR-2003.
XX
XX
PF 28-SEP-2001; 2001WO-US030379.
XX
XX
PR 28-SEP-2001; 2001WO-US030379.
XX
XX
PA (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX
PI Tomich CC, Quinn CL, Arvidson S, Mott JE, Harris DW;
XX
XX
DR WPI; 2003-372002/35.
DR P-PSDB; ADC21340.
XX
XX
PT Identifying an agent that binds a polypeptide and useful for treating or
PT preventing staphylococcal infections, comprises contacting a polypeptide
PT and an agent to form a mixture and determining whether the agent binds
PT the polypeptide.
XX
XX
Example 1; SEQ ID NO 21; 174pp; English.
XX
XX
CC The invention relates to methods of identifying an agent that binds a
CC polypeptide from Staphylococcus aureus by contacting a polypeptide and an
CC agent to form a mixture and determining whether the agent binds the
CC polypeptide. The method is useful in identifying agents that bind gene
CC products critical for the survival of microbes, preferably staphylococcus
CC microbes, including agents that interfere with the function of such gene
CC products. The method is used to identify new agents useful for treating
CC or preventing bacterial infections, particularly those caused by
CC antibiotic-resistant bacteria. This sequence represents the coding region
CC for a protein from Staphylococcus aureus of the invention.
XX
SQ Sequence 837 BP; 314 A; 116 C; 149 G; 258 T; 0 U; 0 Other;

Query Match 5.1%; Score 43.2; DB 9; Length 837;
Best Local Similarity 50.5%; Pred. No. 0.58;
Matches 105; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 403 GGTCAAGCATACAAATCAAGATATGACGAGCAATCTTCTACTTAGGATTATCGCTT 462
Db 820 GGCCAAATCGACCTTGACCATGATGTCGAGAAATACCTGGTGTAGTACATTATATCGT 761
QY 463 CATTTATTAGGAGATGTAATCAGCAATGATGAGCAAACTTTACGAATCTTCTTAT 522
Db 760 AATTATTACATGAAAGCTTCATTTAAATCTTCTTCAATTTATCAATATCAATGA 701
QY 523 CCAATGGGTTTCCATTTCTAAATACGAAATTTTGTGTATACAAATAAAATACTATTT 582
Db 700 CGTTAGCATGCGGACCAATGATTTGATTTCAATCTTCTTATATTAATTCAGCAATTT 641
QY 583 GTTTCAGATAGCAATGGATTTGGAATT 610
Db 640 CCTTTTCAAGTATTGGATAGTAATT 613
```


Db 760 AATTATTACATGAAAGCTTCATTAAAAACCTTCTCAATTTATCAACATATTCATGA 701
 Qy 523 CCAATGGGTTTCCATCTTAATACGAAATTTTGGTTGATACATATAAATAACTATATT 582
 Db 700 CTTAGCATGTGCGACACCAATGGATTGATTTCAATGATCTCTATATAATTCAGCAATTT 641
 Qy 583 GTTTCAGATAGCAATGGATATTGGAAAT 610
 Db 640 CTTTTCAGATATTGGATACATAGATT 613

RESULT 15

AAA70220
 ID AAA70220 standard; DNA; 2700 BP.

XX
 AC AAA70220;

XX 07-NOV-2000 (first entry)

XX Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:353.

XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 KW antimalarial; malaria; protozoa; infection; insecticide; ds.

XX Plasmodium falciparum.

XX WO200025728-A2.

XX 11-MAY-2000.

XX 05-NOV-1999; 99WO-US026796.

XX 05-NOV-1998; 98US-0107131P.

XX (HOFF/) HOFFMAN S.

XX (CARU/) CARUCCI D.

XX (GARD/) GARDNER M.

XX (VENT/) VENTER J C.

PI Hoffman S, Carucci D, Gardner M, Venter JC;

DR WPI; 2000-365347/31.

XX Proteins encoded by chromosome 2 of the human malarial parasite,
 PT Plasmodium falciparum, useful as antimalarial vaccines and in the
 PT diagnosis of P.falciparum infection.

XX Disclosure; Page 542-543; 577pp; English.

XX The present invention describes proteins and their fragments (I) encoded
 CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
 CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
 CC vaccines against P. falciparum infection comprising (I) or (II). (I) and
 CC (II) are useful for the development of vaccines against P. falciparum
 CC infection. (I) and polyclonal antisera or a monoclonal antibody raised to
 CC immunogens comprising the sequences of (I), are useful in the detection
 CC of infection with P. falciparum. Furthermore, (I) (especially when they
 CC are xifins or secreted or membrane proteins) can aid the identification
 CC of drugs to treat or prevent P. falciparum infection, or they can be used
 CC to identify drug resistance in P. falciparum. Sequencing of the
 CC Plasmodium chromosome 2 and the subsequent identification of proteins
 CC encoded by it will help to expand our understanding of parasite biology,
 CC a process hampered by the complexity of the parasitic lifecycle, and
 CC provide new targets for vaccine and drug development. Parasite resistance
 CC to drugs and mosquito resistance to insecticides have led to a resurgence
 CC of malaria in many parts of the world, and there is a pressing need for
 CC vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
 CC represent nucleotide and protein sequences given in the present
 CC invention, but which are not specifically mentioned within the
 CC specification

XX Sequence 2700 BP; 1152 A; 270 C; 390 G; 888 T; 0 U; 0 Other;

Query Match
 Best Local Similarity 5.1%; Score 43.2; DB 3; Length 2700;
 Matches 200; Conservative 0; Mismatches 218; Indels 6; Gaps 2;

Qy 192 TAATACACGATTTGTGAATCCGAATCGAATCGCATTTATTAATAGTGGCGTCTGATTT 251
 Db 294 TTATGATAGGATAAATAATTCTGAAGAAGAAACAAACATTAAACGATGATAATAATGATGA 353
 Qy 252 AGAAAATGGTATTATTCTGCTGATTACGGAATCCTTTATTATGATAATAGTACATATGC 311
 Db 354 TAATAATGGTGAATTTATGATGATGATAATAATAGTGAATGATGATGATGATGATAA 413
 Qy 312 TTCTCACATTTTATGATCGGATACCTG-----GAACAACATATATTCCTTTTTCGGAACAT 366
 Db 414 TAATAACAATGATGATAATAATAATGATGATGATGATGATGATGATGATGATGATGAT 473
 Qy 367 GCAAAAGAAACAGGCGCAAAATATTTTAAACCTTGTCTCAAGCATACCAAAATCAAGAT 426
 Db 474 AAAAGAAAATGATGAATATAAAGATCCTACTTTTTCAGATATCTATATAAAGAAAGCAAAAA 533
 Qy 427 ATGCAGCAAGCATTTCTTACTTAGGATATCGCTTCATTATTTAGGAGATGTGA-ATCA 485
 Db 534 ATGCAATATACGTTGTGAAAAATATTTATGAACAGTTCCGGTTTAATAAGAAGAAATTTAGAAGA 593
 Qy 486 GCCAATGATCGACGCAAACTTTTACGAATCTTTTATCCAATGGTGTTCCTATTTCTAAATA 545
 Db 594 AATAAATGAATCGGATCCCTGAAACAGTTTCAGATAATAGTATGATCTCATCATCAGAAGA 653
 Qy 546 CGAAATTTTGTGATACAAATPAAAAAATAACTATATTTGTTTCAGATAGCAATGGAATTTG 605
 Db 654 AAGTTGTTCTGAGGAATCAGATAAAGAATCAGATAAAGAATCAGATAAAGAATGGAATTTT 713
 Qy 606 GAAT 609
 Db 714 ATAT 717

Search completed: August 2, 2004, 02:22:00
 Job time : 436 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 02:02:34 ; Search time 87 Seconds
(without alignments)
5415.555 Million cell updates/sec

Title: PCT-US03-12556-1

Perfect score: 849
Sequence: 1 atgaaaagaagtattagc.....atcgtatgtaaatcgtaa 849

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	47.4	5.6	7218	1	US-08-232-463-14
C 2	46.6	5.5	3706	4	US-08-913-159-9
C 3	43.4	5.1	9821	4	US-08-956-171E-470
C 4	43.2	5.1	5924	4	US-08-956-171E-130
C 5	42	4.9	1872	4	US-09-543-681A-455
C 6	41.6	4.9	2911	4	US-08-961-527-199
C 7	39.8	4.7	1113	1	US-08-341-538A-1
C 8	39.8	4.7	1113	2	US-08-725-538-1
C 9	39.8	4.7	1353	4	US-09-601-198-37
C 10	39.2	4.6	832	4	US-09-621-976-2813
C 11	39	4.6	640681	4	US-09-790-988-1
C 12	38.4	4.5	352	3	US-08-961-083-67
C 13	38.4	4.5	352	4	US-09-536-784-67
C 14	38.4	4.5	10207	1	US-08-920-822-2
C 15	38.4	4.5	10207	1	US-08-920-827-2
C 16	38.4	4.5	10207	1	US-08-921-177-2
C 17	38.4	4.5	10207	1	US-08-362-577C-2
C 18	38.4	4.5	10207	2	US-08-920-828-2
C 19	38.4	4.5	15766	3	US-09-338-907-73
C 20	38.4	4.5	15766	4	US-09-218-207-73
C 21	38.4	4.5	37950	3	US-09-338-907-183
C 22	38.4	4.5	37950	4	US-09-218-207-183
C 23	38.2	4.5	12172	4	US-09-976-594-125
C 24	38	4.5	1254	4	US-09-134-001C-973
C 25	38	4.5	2649	4	US-09-565-909-1
C 26	38	4.5	5050	4	US-09-565-909-6
C 27	38	4.5	5128	4	US-09-565-909-5

28	37.8	4.5	832	4	US-09-621-976-2813	Sequence 2813, Ap
29	37.6	4.4	426	4	US-09-328-352-2011	Sequence 2011, Ap
30	37.4	4.4	1269	4	US-09-601-198-178	Sequence 178, App
31	37.2	4.4	579	4	US-09-134-001C-1835	Sequence 1835, Ap
32	37	4.4	765	4	US-08-914-479A-3	Sequence 3, Appli
33	37	4.4	855	4	US-08-914-479A-5	Sequence 5, Appli
34	37	4.4	918	3	US-08-937-271-9	Sequence 9, Appli
35	37	4.4	1332	2	US-08-795-475-5	Sequence 5, Appli
36	37	4.4	1617	4	US-09-328-352-2404	Sequence 2404, Ap
37	36.8	4.3	4594	4	US-08-956-171E-200	Sequence 200, App
38	36.8	4.3	7172	4	US-08-961-527-120	Sequence 120, App
39	36.6	4.3	2820	4	US-09-134-000C-2179	Sequence 2179, Ap
40	36.2	4.3	1664976	4	US-08-916-421B-1	Sequence 1, Appli
41	35.6	4.2	740	3	US-09-451-117-1	Sequence 1, Appli
42	35.6	4.2	740	4	US-09-888-655-1	Sequence 1, Appli
43	35.6	4.2	1950	4	US-09-134-001C-1054	Sequence 1054, Ap
44	35.6	4.2	2042	2	US-08-911-434A-2	Sequence 2, Appli
45	35.6	4.2	3006	4	US-09-762-724-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHREIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 5.6%; Score 47.4; DB 1; Length 7218;

```
Best Local Similarity 4.7%; Pred. No. 0.0033;
Matches 12; Conservative 151; Mismatches 92; Indels 0; Gaps 0;

QY 541 AAATACGAAATTTGTTGATACATAAATAAATACTATATGTTTCAGATAGCAATGGA 600
Db 1453 AGATAGAAGAAATTTGTTGATACRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTR 1394
QY 601 TATTGGAATTTGGAAGGACCAACCCAGAGATTGGATTGAAGGAGCAGCGTAGCAGCT 660
Db 1393 RRRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTR 1334
QY 661 AAACAGATTAATCTCGGCTTTGTAACGATACACAAAGATTGGTTTGTAAAGCAGCC 720
Db 1333 RRRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTR 1274
QY 721 GTATCTCAAGATATGCAATAAATGGCGTGGCGGAAGTAACACCGGTACAGGAAGCGT 780
Db 1273 RRRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTRRTTTR 1214
QY 781 TTAATGGAAGCGCAG 795
Db 1213 RRRRTTTRRTTTRRTTTR 1199

RESULT 2
US-08-913-159-9/c
; Sequence 9, Application US/08913159
; Patent No. 6300109
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Plasmid-derived type II
; TITLE OF INVENTION: restriction-modification systems from Lactococcus lactis
; NUMBER OF SEQUENCES: 14
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913,159
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0179/95
; FILING DATE: 17-FEB-1995
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3706 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Lactococcus lactis subsp. cremoris
; STRAIN: W56
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (422..2161)
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /codon_start= 422
; OTHER INFORMATION: /product= "LlaBI methylase"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "orf"
; OTHER INFORMATION: /number= 1
; OTHER INFORMATION: /standard_name= "Gene coding for LlaBI methylase"
; OTHER INFORMATION: /label= m-llaBI
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2464..3360
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /codon_start= 2464
; OTHER INFORMATION: /product= "LlaBI endonuclease"
```

```
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "orf"
; OTHER INFORMATION: /number= 2
; OTHER INFORMATION: /standard_name= "Gene coding for LlaBI
; OTHER INFORMATION: endonuclease"
; OTHER INFORMATION: /label= r-llaBI
US-08-913-159-9

Query Match 5.5%; Score 46.6; DB 4; Length 3706;
Best Local Similarity 47.7%; Pred. No. 0.0041;
Matches 136; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 51 AGTTCAAAGTCTAGTATTGTCACAAACAATAAAGTAGGAAGTCTCGACCGATTTTAAG 110
Db 1010 ATTTAATCGAATAATATATCTTATGAAATATATGAACGCAATGCTGTACCAATCAGCAT 951
QY 111 ATGCTCAGCTGAGGATGAAGCATATAGAGGGGATTAACTCTCATTTGTGGATTGTAATCG 170
Db 950 TGTCTCAATGAAGGAAAAATTCCTTAATCTTTAAATTAATTTATTTGCTATTAAAGACGA 891
QY 171 TGCATTTGACATCATGTCTGTATATACACAGATTGTGAATCCGGAATGAAACTGCAATTAT 230
Db 890 GCTCAGTAAACGCAATGCTGGGAAACCAACCGTTTAAACATTTGGTATGAATATGGCGCATC 831
QY 231 AAATGAGTGGCGTCTGATTTAGAAAAATGGTATTATTCTCTGCTGATTACGAGAAATCCTTA 290
Db 830 TCAAGCATTTGAATAGCCATGTACCTAAATAATTTTCTTACAAATTTCTTTGAATCCAAA 771
QY 291 TTATGATAATAGTACATATGCTTCTCATTCTTATGATCCGGATAC 335
Db 770 CTTTGTATATTTTACAGATTATGCTCTTTTAAATAATGATATGC 726

RESULT 3
US-08-956-171E-470
; Sequence 470, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
```

TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 470:
SEQUENCE CHARACTERISTICS:
LENGTH: 9821 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 470:
US-08-956-171E-470

Query Match 5.1%; Score 43.4; DB 4; Length 9821;
Best Local Similarity 44.3%; Pred. No. 0.045;
Matches 176; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

QY 321 TTATGATCCGGTACTGGAACAACATATATCTCTTTTGGCAACATGCAAAAGAAAGCAGG 380
DB 4223 TGATGGTCAGATTTTCAGAAATGAAAGAAAGAAATTTGAAAGCTTGAATCAAGACG 4282
QY 381 CGCAAAATATTTTAACCTTGTGGTCAAGCATACCAAAATCAAGATATGCGCAAGCAAT 440
DB 4283 TGACATCACTGTTAAAGAAATTTAGTAAGACTGAAAAGAGCAAGACGCGTATTTTAGTAAG 4342
QY 441 CTCTACTTAGGATATCGCTTCATTTTAGGAGATGTGAATCAGCAATGCGCAATGCGCAGC 500
DB 4343 AATGCAAGAAACAGAAATGCTTATTTCAATGACGCAAGCAAGCAATTTAAAGAACG 4402
QY 501 AAACCTTTACGAATCTTCTTATCCAAATGGGTTTCCATTTCTAAATACGAAAATTTTGTGA 560
DB 4403 AGAAAAGCAAGAAAGCAAGAAAGAAAGAGTGGTAAGCAATGAGATGATGTCAT 4462
QY 561 TACAATAAAATAACTATATTTGTTTCAGATAGCAATGGAATTTGGAATTTGGAAGGAGC 620
DB 4463 TGCTATAAAATAACGTCACCTTTCTAAAGTCTGAAAAGATAAATTTGTAGCTATTGC 4522
QY 621 AAACCCAGAGATTCGATTGAAGGAGCAGCGTAGCAGTACCAACAGATTTATCCGCGCT 680
DB 4523 TGATCAAGACATAAGGATGAATGAAGAAAGCAAAATCTAAAGAGATGCTGTAGTAGA 4582
QY 681 TGTGAACGATAGCAAAAGATGTTGTTGTTAAAGCA 717
DB 4583 CGTTGTTAAAGCAAAATAAGATATTGATNAAGAA 4619

RESULT 4
US-08-956-171E-130/c
; Sequence 130, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 130:
SEQUENCE CHARACTERISTICS:
LENGTH: 5924 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 130:
US-08-956-171E-130

Query Match 5.1%; Score 43.2; DB 4; Length 5924;
Best Local Similarity 50.5%; Pred. No. 0.043;
Matches 105; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 403 GGTCAAGCATACCAAAATCAAGATATGCAAGCAAGATTTCTTACTTAGGATATCGCTT 462
DB 5152 GGCCTAATCGACCTTGTACCAAGTATGTGCAAAATAAAGTGGTGTATATATATCGT 5093
QY 463 CATTATTAGGAGATGTGAATCAGCAATGATGCAAGCAATTTTACGAATCTTCTTAT 522
DB 5092 AATTATTCATGAAAGCTTCATTAATAAATCTTCTCAATTTATCAATATCAATGA 5033
QY 523 CCAATGGGTTTCCATTTCTAAATACGAAAATTTTGTGTATACAAATAAAAAATAACTATAT 582
DB 5032 CGTTAGCATGTGCGACCAATGGATTTGATTTTCATGATCTCCTATAAATTCAGCAATTT 4973
QY 593 GTTTCAGATAGCAATGGATATTTGGAAT 610
DB 4972 CTTTTCAGTATTTGGATCTACTMGAAT 4945

RESULT 5
US-09-543-681A-455
; Sequence 455, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 455
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-455

Query Match 4.9%; Score 42; DB 4; Length 1872;
Best Local Similarity 47.7%; Pred. No. 0.057;
Matches 123; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 350 TTCCTTTTGGCAACATGCAAAAGAAAGCAAGCGCGCAATATTTTAACTTCTGCTGTCAG 409
DB 1612 TTACTGCTGTTTGAGAGAAAGATTAAGAAAGAAAGAAAGTATTCACCTTGTGTTTAA 1671
QY 410 CATACCAAAATCAAGATATGCAAGCAATTTCTTACTTAGGATATTCGCTTCAATAT 469
DB 1672 CAGTCATTAAGCAAGATTTCTTACATCATGCTGCTATGCTAGGTTATTTT 1731

QY 470 TAGGAGATGTAATCAGCAATGATCAGCAAACTTTACGAATCTTTCTTATCCAAATGG 529
Db 1732 GATGATATGTTAGTATTCGATACACTGCTTTATCTCCAGATGATTTTAAAGAAAT 1791
QY 530 GTTTCATCTTAAATACGAAATTTTGTGATACAAATAAAATACTATATGTTTTCAG 589
Db 1792 GTCTTAACTGAAGAAATAAAATGTTGCCCATCAATGAAGAGAACAGATATTTAAACAG 1851
QY 590 ATAGCAATGGATATTGA 607
Db 1852 TTAATAGAGAAATAGA 1869

RESULT 6

US-08-961-527-199
; Sequence 199, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961.527

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 199:

SEQUENCE CHARACTERISTICS:

LENGTH: 2911 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-199

Query Match 4.9%; Score 41.6; DB 4; Length 2911;

Best Local Similarity 44.6%; Pred. No. 0.089;

Matches 164; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

QY 212 CGAATGAATCTGATTTTAAATAGTGGCGTGTGATTTAGAAAAATGGTATTTCTG 271
Db 104 CGGAAGAAGATTTTATTTTACCTGTTTATAGGGTGAATTAAGAAAAAGGATACCAATTTG 163
QY 272 CTGATTACGAGATCTTATTATGATATGATCATATGCTTCTCACTTTTATGATCCGG 331
Db 164 ATGGTTGGAAATTTCTGGTTTCGAGGTAAGAAAAAGACGCTGCTATGTTTAAATCTAT 223
QY 332 ATACTGGAAACATATATCTTTTTCGAAACATGCAAAAGAAACAGGCGCAATATTT 391
Db 224 CAAAAGATACCTTTTATAAAACCTGTTTCAAGAAATAGAGGAGAAAAGGAGAGAA 283
QY 392 TTAACCTTGGTGGTCAAGCATACCAAAATCAAGATATGACGCAAGCATTTCTTCTAG 451

Db 284 ATAAACCTACTTTTGATGTATCGAAAAAGAAAGATAACCCCAAGTAACCAATAGTCAAT 343
QY 452 GATTATCGCTTCATTTATTAGGAGATGTGAATCAGCCCAATGCAATGCAAACTTTACGA 511
Db 344 TAAATGAAGTACAGAAAAAGAGATTTTACAAAGAGAAGAGCATTTCACAAAAAATCTGATT 403
QY 512 ATCTTTCTTATCCATGGGTTTCCATCTTAAATACGAAAAATTTTGTTCATACAATAAAA 571
Db 404 CAACTAAGGATGTTACAGCTACAGTCTTGTATAAAAAACAATATCAGTAGTAATCAACTA 463
QY 572 ATAACTAT 579
Db 464 CTAACAAT 471

RESULT 7

US-08-341-538A-1
; Sequence 1, Application US/08341538A
; Patent No. 5817317
; GENERAL INFORMATION:
; APPLICANT: TITBALL DR, RICHARD W
; APPLICANT: WILLIAMSON DR, ETHEL D
; TITLE OF INVENTION: CLOSTRIDIUM PERFRINGENS VACCINES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE
; STREET: 8th FLOOR, 1100 NORTH GLEBBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: UNITED STATES
; ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/341,538A

FILING DATE: 28-NOV-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9210717.6

FILING DATE: 20-MAY-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9215655.3

FILING DATE: 23-JUL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO GB93/01039

FILING DATE: 20-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: CRAWFORD, ARTHUR R

REGISTRATION NUMBER: 25327

REFERENCE/DOCKET NUMBER: 1498-58

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1113 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Clostridium perfringens

STRAIN: double

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1110

FEATURE:

NAME/KEY: CDS

LOCATION: 736..1110

US-08-341-538A-1

	Query Match	4.7%; Score 39.8; DB 1; Length 1113;
	Best Local Similarity	46.8%; Pred. No. 0.18;
	Matches 125; Conservative 0; Mismatches 142; Indels 0; Gaps 0;	
Qy	243 TGC TGATTTAGAAAAATGGTATTTATTCTGCTGATTACGAGAATCCTTATTATGATAATAG 302	
Db	156 TACTTATCCAGATTATGATGAAGATGCATATGATCTATATCAAGATCATTTTCGGATGCC 215	
Qy	303 TACATATGCTTCTCACTTTTATGATCCGGATACCTGGACAACATATATTCCTTTTGCGAA 362	
Db	216 TGATACAGATAATTAATTTCTCAAAGGATAAATGTTGGTATTTAGCTTATCTCTAACCTGA 275	
Qy	363 ACATGCAAAAAGAAAACAGCGCAAAATATTTTAACTTGCTGGTCAAGCATACC AAAATCA 422	
Db	276 CACAGGGGAATCACAAATTAAGAAAAATTTTCAGCATTAGCTAGATATGAATGCCAAAGAGG 335	
Qy	423 AGATATGCAGCAAGCAATCTTCTTACTTAGGATTAATCGCTTCATTTATTAGAGATGTGAA 482	
Db	336 AA ACTATAAACAAGCTACATCTCTATCTTGGAGAGCTATGCATATTTTGGAGATATAGA 395	
Qy	483 TCAGCCCAATGCATGCAGCAAACTTTAC 509	
Db	396 TACTCCCATATCATCTGCTAATGTTAC 422	

RESULT 8

US-08-725-518-1
 ; Sequence 1, Application US/08725518
 ; Patent No. 5851827
 ; GENERAL INFORMATION:
 ; APPLICANT: TITBALL DR, RICHARD W
 ; APPLICANT: WILLIAMSON DR, ETHEL D
 ; TITLE OF INVENTION: CLOSTRIDIUM PERFRINGENS VACCINES
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHYE
 ; STREET: 8th FLOOR, 1100 NORTH GLEBE ROAD
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: UNITED STATES
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/725,518
 ; FILING DATE: 4-OCT-96
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/341538
 ; FILING DATE: 28-NOV-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9210717.6
 ; FILING DATE: 20-MAY-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9215655.3
 ; FILING DATE: 23-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO GB93/01039
 ; FILING DATE: 20-MAY-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CRAWFORD, ARTHUR R
 ; REGISTRATION NUMBER: 25327
 ; REFERENCE/DOCKET NUMBER: 124-536
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1113 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown

RESULT 9

```

US-09-601-198-37/c
; Sequence 37, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-37

```

```

; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Clostridium perfringens
; STRAIN: double
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1110
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 736..1110
US-08-725-518-1

```

RESULT 9

```

US-09-601-198-37/c
; Sequence 37, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-37

```


ATTORNEY/AGENT INFORMATION:
NAME: Brooke, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-67

Query Match 4.5%; Score 38.4; DB 3; Length 352;
Best Local Similarity 44.6%; Pred. No. 0.28;
Matches 150; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 244 GCTGATTAGAAAATGGTATTATTCTGCTGATTACGAGAAATCCTTATTATGATAATAGT 303
DB 8 GGTGAATTAGAAAAGGATACCAATTTGATGGTGGGAAATTTCTGGTTTCGAAGGTAA 67
QY 304 ACATATGCTTCTCACTTTTATGATCCGGATCTGGAACAACATATATTCCTTTTCGGAAA 363
DB 68 AAGACGCTGGCTATGTTTATTAATCTATCAAAAGATACCTTTTATATAAACCTGTATTCAAG 127
QY 364 CATGCAAAAGAAACAGGCGCAAAATATTTTAACTTGTGCTCAAGCATACCAAAATCAA 423
DB 128 AAATAGAGGAGAAAAGAGAGAGAAATAAACCTACTTTTGTATCGAAAAAGAAA 187
QY 424 GATATGACGACGACGATCTTCTTACTTAGGATTTATCGCTTCAATTTATTAGGAGATGTAAT 483
DB 188 GATAACCCACAGTAAACCATAGTCAATTAATGAAGTCACAGAAAAGAGGATTTACAA 247
QY 484 CAGCAATGCTATGACGACAACTTTTACGAATCTTTCTTATCCAAATGGGTTTCCATTTCTAAA 543
DB 248 AGAGAAGAGCATTCACAAAATCTGATTCAACTAAGGATGTTTACAGCTACAGTCTCTGAT 307
QY 544 TACGAAAATTTGTTGATACATAAAAAATTAATAT 579
DB 308 AAAACAATATCAGTAGTAATAATCACTACTACAAAT 343

RESULT 13
US-09-536-784-67
Sequence 67, Application US/09536784
Patent No. 6573082
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/536,784
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997

ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-536-784-67

Query Match 4.5%; Score 38.4; DB 4; Length 352;
Best Local Similarity 44.6%; Pred. No. 0.28;
Matches 150; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 244 GCTGATTAGAAAATGGTATTATTCTGCTGATTACGAGAAATCCTTATTATGATAATAGT 303
DB 8 GGTGAATTAGAAAAGGATACCAATTTGATGGTGGGAAATTTCTGGTTTCGAAGGTAA 67
QY 304 ACATATGCTTCTCACTTTTATGATCCGGATCTGGAACAACATATATTCCTTTTCGGAAA 363
DB 68 AAGACGCTGGCTATGTTTATTAATCTATCAAAAGATACCTTTTATATAAACCTGTATTCAAG 127
QY 364 CATGCAAAAGAAACAGGCGCAAAATATTTTAACTTGTGCTCAAGCATACCAAAATCAA 423
DB 128 AAATAGAGGAGAAAAGAGAGAGAAATAAACCTACTTTTGTATCGAAAAAGAAA 187
QY 424 GATATGACGACGACGATCTTCTTACTTAGGATTTATCGCTTCAATTTATTAGGAGATGTAAT 483
DB 188 GATAACCCACAGTAAACCATAGTCAATTAATGAAGTCACAGAAAAGAGGATTTACAA 247
QY 484 CAGCAATGCTATGACGACAACTTTTACGAATCTTTCTTATCCAAATGGGTTTCCATTTCTAAA 543
DB 248 AGAGAAGAGCATTCACAAAATCTGATTCAACTAAGGATGTTTACAGCTACAGTCTCTGAT 307
QY 544 TACGAAAATTTGTTGATACATAAAAAATTAATAT 579
DB 308 AAAACAATATCAGTAGTAATAATCACTACTACAAAT 343

RESULT 14
US-08-920-812-2
Sequence 2, Application US/08920812
Patent No. 5763188
GENERAL INFORMATION:
APPLICANT: Ohno, Teuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,812
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```
/ APPLICATION NUMBER: US 08/362,577
/ FILING DATE: 27-MAR-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rin-Laures, Li-Hsien
/ REGISTRATION NUMBER: 33,547
/ REFERENCE/DOCKET NUMBER: 19036/32420
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/474-6300
/ TELEFAX: 312/474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 10207 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: Genomic DNA
/ ORIGINAL SOURCE:
/ ORGANISM: Staphylococcus aureus
/ STRAIN: Clinical Isolate SA-24
/ US-08-920-812-2

Query Match      4.58; Score 38.4; DB 1; Length 10207;
Best Local Similarity 50.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 204 TGTGAATCCGAATGAAACTGCATTATTAATGAGTCGCGTGTGATTTAGAAAATGGTAT 263
Db 3818 TGCTAAGCAACCAAAATACATTATTGAGCAATTCGCTGTGTTAGATGATGAAT 3877

QY 264 TTATTCGCTGATACAGAAATCCTTATTATGATAATAGTACATATGCTTCTCATTTTA 323
Db 3878 TGTGCTGATAGTTGGTTAGATGCTTATCAAGTAATTAGATAGCCATCATTTAAATCA 3937

QY 324 TGATCCGGATACCTGGAACAACATATATCCTTTTGGCAACATGCAAAAGAACAGGCGC 383
Db 3938 AGGTTAGATTTATTAATGTCAGCAATTAACGTTTGACAATGAAACTGTAAAATTAGGTGA 3997

QY 384 AAAATATTTTAA 395
Db 3998 AACGTTGCTTAA 4009
```

```
RESULT 15
US-08-920-827-2
/ Sequence 2, Application US/08920827
/ Patent No. 5770375
/ GENERAL INFORMATION:
/ APPLICANT: Ohno, Tsuneva
/ APPLICANT: Matsuhisa, Akio
/ APPLICANT: Uehara, Hirotsugu
/ APPLICANT: Eda, Soji
/ TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
/ STREET: 6300 Sears Tower, 233 South Wacker Drive
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: United States of America
/ ZIP: 60606-6402
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/920,827
/ FILING DATE: 29-AUG-1997
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/362,577
/ FILING DATE: 27-MAR-1995
```

```
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rin-Laures, Li-Hsien
/ REGISTRATION NUMBER: 33,547
/ REFERENCE/DOCKET NUMBER: 19036/32420
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/474-6300
/ TELEFAX: 312/474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 10207 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: Genomic DNA
/ ORIGINAL SOURCE:
/ ORGANISM: Staphylococcus aureus
/ STRAIN: Clinical Isolate SA-24
/ US-08-920-827-2

Query Match      4.5%; Score 38.4; DB 1; Length 10207;
Best Local Similarity 50.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 204 TGTGAATCCGAATGAAACTGCATTATTAATGAGTCGCGTGTGATTTAGAAAATGGTAT 263
Db 3818 TGCTAAGCAACCAAAATACATTATTGAGCAATTCGCTGTGTTAGATGATGAAT 3877

QY 264 TTATTCGCTGATACAGAAATCCTTATTATGATAATAGTACATATGCTTCTCATTTTA 323
Db 3878 TGTGCTGATAGTTGGTTAGATGCTTATCAAGTAATTAGATAGCCATCATTTAAATCA 3937

QY 324 TGATCCGGATACCTGGAACAACATATATCCTTTTGGCAAAACATGCAAAAGAACAGGCGC 383
Db 3938 AGGTTAGATTTATTAATGTCAGCAATTAACGTTTGACAATGAAACTGTAAAATTAGGTGA 3997

QY 384 AAAATATTTTAA 395
Db 3998 AACGTTGCTTAA 4009
```

Search completed: August 2, 2004, 04:12:30
Job time : 102 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 01:48:00 ; Search time 2814 Seconds

(without alignments)
9009.592 Million cell updates/sec

Title: PCT-US03-12556-1

Perfect score: 849

Sequence: 1 atgaaaagaagattagc.....atacgtatgtaaatcgctaa 849

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_nam:*

23: em_gss_nus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
c 1	52.2	6.1	994	13 BX414650	BX414650 BX414650
2	51.8	6.1	844	29 BX139987	BX139987 Danio rer
3	51.4	6.1	605	12 BM162501	BM162501 EST565024
c 4	50.4	5.9	1101	29 CNS0039G	AL063921 Drosophil

c 5	50	5.9	977	28	AZ687764	ENTHD89TF
6	49.2	5.8	634	12	BM164871	EST567394
c 7	48.2	5.7	776	28	BZ063272	1k269h12.
c 8	48	5.7	754	12	BX376728	BX376728
9	47.8	5.6	1201	13	BX355654	BX355654
10	47	5.5	862	28	AZ677978	ENTHJ38TR
11	47	5.5	1101	29	CNS0172Q	AL108704 Drosophil
c 12	46.8	5.5	1190	29	CNS020N7	AL206908 Tetraodon
13	46.4	5.5	686	29	CNS03412	AL226880 Tetraodon
c 14	46	5.4	1201	13	BX361152	BX361152
15	46	5.4	1201	13	BX462896	BX462896
16	45	5.3	357	12	BI814749	PFEST0aa0
17	45	5.3	1175	28	BZ696793	SP_Ba008
c 18	44.8	5.3	1201	9	AL536104	AL536104
19	44.6	5.3	404	28	AZ140343	AZ140343 SP_0001 A
20	44.6	5.3	740	29	CNS02YR2	AL220040 Tetraodon
21	44.6	5.3	843	29	CNS00CS1	AL059666 Drosophil
c 22	44.4	5.2	541	29	CNS04KUS	AL295309 Tetraodon
23	44.2	5.2	833	28	BH154104	BH154104 ENTRBS8TF
c 24	44.2	5.2	1101	29	CNS00LT2	AL078714 Drosophil
c 25	44	5.2	613	12	CF181891	CF181891 ISO2H2F I
26	44	5.2	666	14	BJ401356	BJ401356
c 27	44	5.2	829	29	CNS00A4K	AL054917 Drosophil
c 28	44	5.2	838	28	AZ195507	AZ195507 SP_1030 A
c 29	44	5.2	1056	13	BX415058	BX415058
c 30	44	5.2	1124	13	BX436282	BX436282
31	44	5.2	1200	13	BX437758	BX437758
c 32	43.8	5.2	575	28	BZ780846	BZ780846 i120h11.g
c 33	43.8	5.2	705	12	BJ446775	BJ446775
34	43.8	5.2	1201	13	BX461310	BX461310
c 35	43.6	5.1	429	9	AU262155	AU262155
c 36	43.6	5.1	494	28	AZ179729	AZ179729 SP_0166 A
c 37	43.6	5.1	526	9	AU037906	AU037906
38	43.6	5.1	543	29	BX233514	BX233514 Danio rer
39	43.6	5.1	816	12	BJ437362	BJ437362
40	43.6	5.1	859	29	CNS00KLL	AL077728 Drosophil
41	43.6	5.1	992	29	CNS06P6W	AL408990 T7 end of
c 42	43.4	5.1	1201	9	AL536104	AL536104
c 43	43.2	5.1	470	9	AU034262	AU034262
c 44	43.2	5.1	609	12	BJ444006	BJ444006
c 45	43.2	5.1	1101	29	CNS016LI	AL106896 Drosophil

ALIGNMENTS

RESULT 1

BX414650/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX414650 Homo sapiens THYMUS cDNA clone CS0CAP001YN02
3-PRIME, mRNA sequence.
BX414650.1 GI:30763455
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 994)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6015.f
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitroGen/Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0CAP001D001NP1.
Location/Qualifiers
1 . 994

FEATURES

source

Db 800 KXADDDKDAADRRDAATTTTTRTDDDKWKTDTWRAADRTWDRDDDDDR 741
Qy 503 ACTTACGAATCTTTCTTATCAATGGGTTTCCATCTTAATACGAAATTTTGTGATA 562
Db 740 AGTAGKWRRTWKRRKRDTRWDADDDTARDRRRRGGDADAGKGGKTKGRRRRD 681
Qy 563 CAATATAAATAAATCAATATTTGTTTCAGATAGCAATGGATATTTGGAATTTGGAAGAGCA 622
Db 680 RATWDRDADWADAATTTTDDDDDKRRRRKRGARRRTTARAANDWWTWKANDWA 621
Qy 623 ACCCAGAAGATTGGATTGAAGAGAGCGGTTAGCAGCTTAAACAAGATTATCTCGCGTTG 682
Db 620 KWDWKTADRWDRMAADTTWDKARDWAKARAWARRDRARAARDRTWTGKTITTA 561
Qy 683 TGAACGATACGACAAAGATTGGTTT 708
Db 560 TWTWAAABAANWAWAATTTATWTT 535

RESULT 5
AZ687764/c
LOCUS
DEFINITION
ENTHD89TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
genomic, genomic survey sequence.
ACCESSION
AZ687764
VERSION
AZ687764.1 GI:11824910
KEYWORDS
GSS.
SOURCE
Entamoeba histolytica
ORGANISM
Entamoeba histolytica
REFERENCE
1 (bases 1 to 977)
Lofus, B., Van Aken, S., and Fraser, C.
Determination of clone end sequences from Entamoeba histolytica
HM1:IMSS sheared DNA library
Unpublished (2000)
JOURNAL
COMMENT
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjoftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 58
High quality sequence stop: 493.
Location/Qualifiers
1..977
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHO31; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark, C.G.,
and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

FEATURES
source
Query Match 5.9%; Score 50; DB 28; Length 977;
Best Local Similarity 49.4%; Pred. No. 0.084;
Matches 157; Conservative 0; Mismatches 160; Indels 1; Gaps 1;

Qy 375 AACAGGCGCAAAATATTTTAACTCTGCTCAAGCATACCAAAATCAAGATATGACGA 434
Db 560 ATCAAGTGCAAAAGATTTTGACATGAATGAAGGGTTAAACAAATAACGTTTACACCA 501
Qy 435 AGCAATCTTCTACTTAGGATTAATCGCTTCATTA-TTTAGGAGATGTGAATCAGCAATGC 493
Db 500 TTGAAGAAGTTTATAAATATGCAAAATTCATTAGTGGAAACGGTTGCACAAACCTCCTT 441
Qy 494 ATCAGCAAACTTTAGCAATCTTTTATCCAAATGGGTTTCCATCTTAATACGAAAT 553
Db 440 ATCAAAATAGCGCTTCATCTCTTATCCATATCATATTCCTGTATGTACCACTG 381
Qy 554 TTGTTGATACATATAAATAAATACTATATTTGTTTTCAGATAGCAATGATATTTGAAATGGA 613
Db 380 GTGTTGATTTATTAGAAAATGCTGGATGGAATAGAGATATTACTGGGTACAGTATTCAG 321
Qy 614 AAGGAGCAAAACCCAGAAGATTGGATTGAAGAGCAGCGGTAGCAGCTTAACAAGATTATC 673
Db 320 ATGAAGTAAAGAAGACAAACCAAGAAAGAACCTGAGTTTGAAGTTGATATGATAATG 261
Qy 674 CTGCGGTTGTGAACGATA 691
Db 260 ATGTTGTTCTTATGGA 243

RESULT 6
BM164871
LOCUS
DEFINITION
EST567394 PyBS Plasmodium yoelii yoelii cDNA clone PYCMD07 5' end,
mRNA sequence.
ACCESSION
BM164871
VERSION
BM164871.1 GI:17310552
KEYWORDS
EST.
SOURCE
Plasmodium yoelii yoelii
ORGANISM
Plasmodium yoelii yoelii
REFERENCE
1 (bases 1 to 634)
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
AUTHORS
Carlton, J.M., Daly, T.M., Long, C.A., Bergman, L.W., Vaidya, A.B.,
Fraser, C.M. and Carucci, D.J.
Plasmodium yoelii EST project at TIGR
Unpublished (2001)
TITLE
JOURNAL
COMMENT
Contact: Jane Carlton
Parasite Genomics Group
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-530-9319
Fax: 301-538-0208
Email: carlton@tigr.org
For clone info, please contact the Malaria Research and Reference
Reagent Resource Center, ATCC
http://www.malaria.mr4.org/mr4pages/index.html
Seq primer: ADF.

FEATURES
source
1..634
/organism="Plasmodium yoelii yoelii"
/mol_type="mRNA"
/strain="17XL"
/sub_species="yoelii"
/db_xref="taxon:73239"
/clone="PYCMD07"
/dev_stage="Asexual blood stages"
/lab_host="E. coli XL-1 Blue"
/note="Vector: pAD-GAL4; At 20-25% parasitemia, blood was
collected from BALB/cByJ mice infected with Py17XL
parasites, and leukocytes removed by passage over
microcrystalline cellulose columns. Total RNA was
isolated using the guanidium isothiocyanate method, and
mRNA isolated using oligo(dT)-cellulose chromatography.
First strand cDNA synthesis was completed using a 50-base
primer and reverse transcriptase in the presence of
5-methyl dCTP. After second strand synthesis, uneven
Location/Qualifiers

termini were treated with Pfu DNA polymerase and EcoRI adaptors ligated to the blunt ends. The sample was cleaved with XhoI and separated on a Sephacryl S-500 column. Size-fractionated cDNA was precipitated and ligated to HybriZAP arms directionally using EcoRI-XhoI cleaved arms. After packaging, the phagemid vector (pAD-GAL4) was excised from the HybriZAP vector and plasmid DNA isolated."

ORIGIN

Query Match 5.8%; Score 49.2; DB 12; Length 634;
Best Local Similarity 50.9%; Pred. No. 0.12;
Matches 117; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 547 GAAATTTTGTGATACATAAAAAATACTATATATTTCTTCAGATGAGCAATGGATTTGG 606
DB 95 GAAAGTTTTTGGAGGAAAAATAAAAAAGATAAAATGGGTATGAGAAAAATGAAAAAGTGA 154
QY 607 AATTGGAAGGACCAACCCAGAGATTGGATTGAAGGAGCAGCGGTAGCAGCTAAACAA 666
DB 155 GATAGTACAGAGAAAAGTCAGATGAAGAAGATGAAGAGGATGAAGAGATGAAGAGAA 214
QY 667 GATTATCTCGGCTGTGACATACGACAAAGATTGTTTGTAAAAAGCAGCCGTATCT 726
DB 215 GATGAGATGAAGAGATGAAGAGAGAAAGATGATCTCTGAAAAATTAAGAAATTAAT 274
QY 727 CAAGAATATGAGATAAATGGCGTGGGAAGTAACACCGGTGACAGGAAA 776
DB 275 AAAGAAGAACTATAGATGTCCACCATTAGCTCAATATAGAGGAGA 324

RESULT 7

BZ063272/c
LOCUS lk269hl2.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
DEFINITION sequence.

ACCESSION

BZ063272.1 GI:23673911

VERSION

KEYWORDS

SOURCE

ORGANISM

Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE

1 (bases 1 to 776)
Delehaunty, K., Fewell, G., Fulton, L., McComb, W.R., Miner, T.,

AUTHORS

Nash, W., Rabinowicz, P.D. and Wilson, R.K.

TITLE

Whole genome shotgun reads from Brassica oleracea

JOURNAL

COMMENT

Unpublished (2002)
Contact: Richard K. Wilson

Genome Sequencing Center
Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: lk269 row: h column: 12

Seq primer: -21UpPOT forward

Class: shotgun

High quality sequence start: 8

High quality sequence stop: 551.

FEATURES

source

1..776
/organism="Brassica oleracea"

/mol_type="genomic DNA"

/db_xref="taxon:3712"

/clone_lib="B.oleracea002"

/note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T0100DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Query Match 5.7%; Score 48.2; DB 28; Length 776;
Best Local Similarity 50.7%; Pred. No. 0.23;
Matches 113; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 268 TCTGCTGATTACGAGAATCCCTTATTATGATAATAGTACATATGCTTCTCAGCTTTTATGAT 327
DB 750 TCTTATTATGTCGAGAACAGTATTTCTTATTATCAAGTTGTATCTTATTTATTATTAN 691
QY 328 CCGGATACCTGGAAACAACATATATTCCTTTTTCGAAAAACATGCAAAAGAAACAGGCCAAAA 387
DB 690 AATTATACCTGTTCTGGAGTACATAAAGTTCTTAAAGAAATAAAAAATAAATGAAGTAAAT 631
QY 388 TATTTTAACTTGTGCTCAAGCATACCAAAATCAAGATATGCAGCAGGATTTCTTCTTAC 447
DB 630 AATATATGCTCGGAAAGAAAAATAGAAAAATAAATAAATTTGAAATATACAAATNTAGTG 571
QY 448 TTAGGATTATCGCTTTCATTATTTTAGGAGATGTGAATCAGCCAA 490
DB 570 GGGTATTAGGACTCTAATTTTGAAGAGGCCCATTTATTGAA 528

RESULT 8
BJ376728/c
LOCUS BJ376728 Dictyostelium discoideum cDNA library, CF Dictyostelium
DEFINITION discoideum cDNA clone ddc29o13 3', mRNA sequence.

ACCESSION BJ376728.1 GI:19286111
VERSION BJ376728
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE 1 (bases 1 to 754)
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the culmination stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tehin@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
source
1..754
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddc29o13"
/sex="mat A"
/dev_stage="Culmination stage"
/clone_lib="Dictyostelium discoideum cDNA library, CF"

ORIGIN

Query Match 5.7%; Score 48; DB 12; Length 754;
Best Local Similarity 45.7%; Pred. No. 0.26;
Matches 168; Conservative 0; Mismatches 200; Indels 0; Gaps 0;

QY 193 AATCAACAGATTGGAATCCGAATGAAACTGCATTATTAATAGTGGCGTGTGATTATA 252
DB 487 AATAAAATGGGTGATGATTATGATGAAGAGAAATTTAATAAGAAATTTAATAAGAAATTA 428
QY 253 GAAATGGTATTATTCGCTGATTACGAGAAATCCTTATTATGATTAATAGTACATAGCT 312
DB 427 GATATTACAGTAGTAGTGAAGAATATGATGATTAATTAATAATAAACAATGATGAT 368
QY 313 TCTCAGCTTTTATGATCCGATACCTGGAACAACATATATTCCTTTTGGAAAAACATGCAAAA 372
DB 367 GATGATGATATGATGAAGATGATGCTGGAATATTATCAAGAAGAGGATTTGAAGAT 308
QY 373 GAAACAGCGCCAAAAATATTTTAACTTGGTGGTCAAGCATACCAAAATCAAGATATGAG 432

```

Db      307 GAAAAAAGCTGAAATTCABATTAATAAATAATTAAGAGGATGATCATGATGATGAT 248
QY      433 CAAGCATCTTCTACTTACCTAGATATTCCTGATATTTAGGAGATGATGATGATGATGAT 492
Db      247 GATGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 188
QY      493 CATGACGAAATCTTACGATCTTCTTCTATCCATGCTGCTTCTTCTTCTTCTTCTTCTTCT 552
Db      187 GTTGATATAATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 128
QY      553 TTGCTTGA 560
Db      127 GATAATGA 120

```

```

RESULT 9
LOCUS   BX355654
DEFINITION BX355654 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI002YE12 3-PRIME, mRNA sequence.
ACCESSION BX355654
VERSION   BX355654.1 GI:30371982
SOURCE   EST.
ORGANISM Homo sapiens (human)

```

```

REFERENCE
AUTHORS  Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE    Full-length cDNA libraries and normalization
JOURNAL  Unpublished (2001)
COMMENT  Genoscope - Centre National de Sequencage

```

```

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI002BC06NP1.
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI002YE12"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

```

FEATURES
source

```

```

ORIGIN

```

```

Query Match      5.6%; Score 47.8; DB 13; Length 1201;
Best Local Similarity 28.6%; Pred No. 0.31;
Matches 131; Conservative 110; Mismatches 214; Indels 3; Gaps 1;

QY      385 AAATATTTAACTTGGTGGTCAAGCATACCAAAATCAAGATATGACGACGATTTCTTC 444
Db      523 ATATATTTTATTTTATTAATTTGGTAAATAAATAAATTTAATAATAAATAAATAA 582
QY      445 TACTTAGGATATCGCTTCATTTAGGAGATGATGATGATGATGATGATGATGATGATGAT 504
Db      593 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 642
QY      505 TTTAGGAATCTTCTTCTTCAATGGGTTTCTTCAATCTTCAATACGAAATTTTGTGTACA 564
Db      643 DAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 702
QY      565 ATAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 624
Db      703 WAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 759

```

```

QY      625 CCAGAAGATTGGATTGAAGGAGCAGCGGTAGCAGTAAACAAGATTATCTCTGGCGTGTG 684
Db      760 AAKKAAKAAKGAADKGAADBDWKGAADKGAADKGAADKGAADKGAADKGAADKGAADKGA 819
QY      685 AACGATACGACAAAGATGTTGTTGTAAGACGCGGTATCTCAAGATATGCGAGATAAA 744
Db      820 GKADGAAAKGGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 879
QY      745 TGGCGTGGCGGAAGTAACACCGGTGACAGGAAGCGTTTAAATGGAAGCGCGCGGTACA 804
Db      880 KGKGGKKKKKAAADKGAADKGAADKGAADKGAADKGAADKGAADKGAADKGAADKGA 939
QY      805 GCTGTTATATTCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 842
Db      940 ARADKGAAGATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 977

```

```

RESULT 10
LOCUS   AZ677978
DEFINITION ENTHJ38TR Entamoeba histolytica Sheared DNA linear GSS 14-DEC-2000
genomic, genomic survey sequence.
ACCESSION AZ677978
VERSION   AZ677978.1 GI:11815244
KEYWORDS  GSS.
SOURCE    Entamoeba histolytica
ORGANISM  Eukaryota; Entamoebidae; Entamoeba.

```

```

REFERENCE
AUTHORS  Loftus, B., Van Aken, S. and Fraser, C.
TITLE    Determination of clone end sequences from Entamoeba histolytica
JOURNAL  HMI:IMSS sheared DNA library
COMMENT  Unpublished (2000)

```

```

Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 15
High quality sequence stop: 862.
Location/Qualifiers
1. .862
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHS01; Site 1: Bat I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."
```

```

FEATURES
source

```

```

ORIGIN

```

```

Query Match      5.5%; Score 47; DB 28; Length 862;
Best Local Similarity 48.9%; Pred. No. 0.46;
Matches 154; Conservative 0; Mismatches 160; Indels 1; Gaps 1;

```


784	KASWAAMCA	YACKYRAAAAAWAHWMTKTYCTTTTMTAYMYMTGRITWTAFAAMKRM	843
Qy	345	ATATATTCCCTTTTGCAGAAACATGCACAAAGAAACAGGGCGCAAAATATTTTAACTTGTCTGG	404
Db	844	DNKAMKATVTATAWAWAWATTATATTTTAWWATWTATWMTKTYTYTBTBTATWTAT	903
Qy	405	TCAAGCATACAAAATCAAGATATGCAGCAAGCATCTTCTACTTAGGATTATCGCTTCA	464
Db	904	ATATTATWATWATATATATTTWMTWAAWAAATTTWMMWTATWTATATATATATATATATW	963
Qy	465	TTATTTAGGAGATGTGAATCAGCCAATGCATGCAGCAAACTTTTACGGAATCTTCTCTTATCC	524
Db	964	ATATYWATAWATATATWATAAATAWAAATATATWMTATATATATATATATATATATATATAT	1023
Qy	525	AATGGGTTTCCATCTTCAAATACGAAAATTTTGTGTGATACAAATAAAAAATAACTATATGT	584
Db	1024	WATATWTTATATAATATATATATATWATWTATWMTTWTATATATATATATATATATATATW	1083
Qy	585	TTCCAGATAGCAATGGATA	602
Db	1084	WTWMTWATWMTWATAWATA	1101

RESULT 12
CNS020N7/c
LOCUS
DEFINITION
1190 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
153P04 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL206908
AL206908.1 GI:7865727
GSS; genome survey sequence.
Tetraodon nigroviridis
Tetraodon nigroviridis
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.

AUTHORS Roest Crolius, H., Jallou, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fzames, C., Wncker, P., Brotier, P., Quetier, F., Saurin, W. and Weissenbach, J.

TITLE Estimate of human gene number provided by genome-wide analysis using Tetraodon nigriviridis DNA sequence

JOURNAL Nat. Genet. 25 (2), 235-238 (2000)

TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish *Tetraodon nigroviridis*

JOURNAL Genome Res 10 (7) 939-949 (2000)

```

COMMENT
- Web : www.genoscope.cns.fr}
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

FEATURES
    source
        1. .1190
            Location/Qualifiers
                /organism="Tetraodon nigroviridis"
                /mol_type="genomic DNA"
                /db_xref="taxon:99883"
                /clone="153P04"

```



```
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODM001YL14"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN
Query Match      5.4%; Score 46; DB 13; Length 1201;
Best Local Similarity 36.7%; Pred. No. 0.87;
Matches 127; Conservative 37; Mismatches 182; Indels 0; Gaps 0;

Qy 226 TTATTAAATGAGTGGCGTCTGATTTAGAAATGATTTATCTGCTGATTCAGAGAT 285
Db 954 TTTTWWAWTTKKGGKTTTWWTTTTTAAAAAAAAMWTTTTTTTTTWWAAAAAAWT 895

Qy 286 CCTATTATGATAATAGTACATATCTCTCACTTTTATGATCCGGATCTGGAACAACA 345
Db 894 TTTTTTTTTTTRARKTTTTTAAAWTTTTTTTTTTTTTTTTTTTTTTTAAATAATWTTGAA 835

Qy 346 TATATTCCTTTTGGCAACATGCAAAAGAAACAGCGCAAAATATTTTAACTTCTGCTGGT 405
Db 834 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATTT 775

Qy 406 CAAGCATACCAAAATCAAGATATGACGACAGATCTTCTACTTAGGATTTATCGCTTCAT 465
Db 774 AWWWWWTTWWAAWATAAAWAAWATAAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 715

Qy 466 TATTAGGAGATGTGAATCAGCAATGATGACGCAAACTTTTACGAATCTTCTTATCCA 525
Db 714 TTTTAAARRANNNTNNNNAANAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 655

Qy 526 ATGGGTTTCCATCTTAAATACGAAATTTTGTGTATACAAATAAAAA 571
Db 654 TTTTAAATGTAAGTAAGTTTATTAAATTAATTAAGATCTTGAAAA 609

RESULT 15
BX462896
LOCUS
DEFINITION
BX462896 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CSODM001YL14 3-PRIME, mRNA sequence.
BX462896
VERSION
KEYWORDS
SOURCE
EST.
GR:31019600
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4458.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODM001DF07NP1&cluster=4458.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODM001DF07NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODM001YL14"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN
Query Match      5.4%; Score 46; DB 13; Length 1201;
Best Local Similarity 39.1%; Pred. No. 0.87;
Matches 148; Conservative 17; Mismatches 214; Indels 0; Gaps 0;

Qy 225 ATTATTAATGAGTGGCGTCTGATTTAGAAATGATTTATCTGCTGATTCAGAGAA 284
Db 699 AATAWNAATAAATTTATTANTATATTTTATAWAATANWATTTTATTANANANAAAAATAA 758

Qy 285 TCCTTATTATGATATAGTACATATGCTTCTCACTTTTATGATCCGGATCTGGAACAAC 344
Db 759 ATAAATTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 818

Qy 345 ATATATTCCTTTTGGCAACATGCAAAAGAAACAGCGCAAAATATTTTAACTTCTGCTGG 404
Db 819 TATAATTTATATATAATWAAAAAATAATTTATATATATAATAATAATAATAATAATN 878

Qy 405 TCAAGCATACCAAAATCAAGATATGACGACAGATCTTCTACTTAGGATTTATCGCTTCA 464
Db 879 TAATTATAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATN 938

Qy 465 TTATTTAGGAGATGTGAATCAGCAATGATGACGCAAACTTTACGAATCTTCTTATCC 524
Db 939 TTATTTAAATTAATAAANWNTAATAATAATAATAATAATAATAATAATAATAATAATTA 998

Qy 525 AATGGTTTCCATCTTAAATACGAAATTTTGTGTATACAAATAAAAAATACTATATGT 584
Db 999 ATTAATAWTAATAWNTAATAATAATAATAATAATAATAATAATAATAATAATAATA 1058

Qy 585 TTCAGATAGCAATGGATAT 603
Db 1059 WATATANANTATAATAAT 1077

Search completed: August 2, 2004, 04:10:48
Job time : 2820 secs
```

This Page Blank (uspto)